

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION  
International Bureau



INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification <sup>6</sup> : <b>C07K 14/00</b>		A2	(11) International Publication Number: <b>WO 99/33868</b>
			(43) International Publication Date: <b>8 July 1999 (08.07.99)</b>
<p>(21) International Application Number: <b>PCT/EP98/08563</b></p> <p>(22) International Filing Date: <b>18 December 1998 (18.12.98)</b></p> <p>(30) Priority Data: <b>9727262.9 24 December 1997 (24.12.97) GB</b></p> <p>(71) Applicant (<i>for all designated States except US</i>): <b>SMITHKLINE BEECHAM BIOLOGICALS S.A. [BE/BE]; Rue de l'Institut 89, B-1330 Rixensart (BE).</b></p> <p>(72) Inventors; and</p> <p>(75) Inventors/Applicants (<i>for US only</i>): <b>DALEMANS, Wilfried, L., J. [BE/BE]; SmithKline Beecham Biologicals S.A., Rue de l'Institut 89, B-1330 Rixensart (BE). GERARD, Catherine, Marie, Ghislaine [BE/BE]; SmithKline Beecham Biologicals S.A., Rue de l'Institut 89, B-1330 Rixensart (BE).</b></p> <p>(74) Agent: <b>DALTON, Marcus, Jonathan, William; SmithKline Beecham, Two New Horizons Court, Brentford, Middlesex TW8 9EP (GB).</b></p>			
<p>(81) Designated States: <b>AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIGO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).</b></p> <p><b>Published</b> <i>Without international search report and to be republished upon receipt of that report.</i></p>			
<p>(54) Title: <b>VACCINE</b></p> <p>(57) Abstract</p> <p>The present invention provides Human Papilloma Virus (HPV) fusion proteins, linked to an immunological fusion partner that provides T helper epitopes to the HPV antigen. Vaccine formulations are provided that are useful in the treatment or Prophylaxis of HPV induced tumours.</p>			

**FOR THE PURPOSES OF INFORMATION ONLY**

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakhstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LJ	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		

## VACCINE

The present invention relates to vaccine compositions, comprising an E6 or/ and E7 or E6, E7 fusion protein from an HPV strain optionally linked with an 5 immunological fusion partner and formulated with a CpG containing oligonucleotide into vaccines that find utility in the treatment or prophylaxis of human papilloma virus induced tumours or lesions. In particular, the present invention relates to vaccines comprising fusions proteins, comprising a protein or part of a protein that provides T helper epitopes (such as protein D from Haemophilus influenzae B) and an antigen 10 from a human-papilloma virus (eg comprising an E6 or E7 protein from HPV 16 or 18 strain associated with cancer) that find utility in the treatment or prophylaxis of human papilloma induced tumours, wherein the vaccine is formulated with a CpG containing oligonucleotide as an adjuvant.

Papillomaviruses are small naked DNA tumour viruses (7.9 kilobases, double 15 strand), which are highly species-specific. Over 70 individual human papillomavirus (HPV) genotypes have been described. Papillomaviruses are classified on the basis of species of origin (human, bovine etc.) and of the degree of genetic relatedness with other papillomaviruses from the same species. HPVs are generally specific for the skin or mucosal surfaces and have been broadly classified into "low" and "high" risk 20 viruses.

Low risk HPVs usually cause benign *lesions* (warts or papillomas) that persist for several months or years. High risk HPVs are associated with pre-neoplastic lesions and cancer. The strongest positive association between an HPV virus and 25 human cancer is that which exist between HPV 16 and 18 and cervical carcinoma. More than ten other HPV types have also been found in cervical carcinomas including HPV 31 and HPV 33 although at less frequency.

Genital HPV infection in young sexually active women is common and most individuals either clear the infection, or if lesions develop, these regress. Only a subset of infected individuals has lesions which progress to high grade intraepithelial 30 neoplasia and only a fraction of these progress further to invasive carcinoma.

The molecular events leading to HPV infection have not been clearly established. The lack of an adequate *in vitro* system to propagate human

papillomaviruses has hampered the progress to a best information about the viral cycle.

Today, the different types of HPVs have been isolated and characterised with the help of cloning systems in bacteria and more recently by PCR amplification. The 5 molecular organisation of the HPV genomes has been defined on a comparative basis with that of the well characterised bovine papillomavirus type 1 (BPV1).

Although minor variations do occur, all HPVs genomes described have at least seven early genes, E1 to E7 and two late genes L1 and L2. In addition, an upstream regulatory region harbors the regulatory sequences which appears to control most 10 transcriptional events of the HPV genome.

E1 and E2 genes are involved in viral replication and transcriptional control, respectively and tend to be disrupted by viral integration. E6 and E7 are involved in viral transformation. E5 has also been implicated in this process.

In the HPVs involved in cervical carcinoma such as HPV 16 and 18, the 15 oncogenic process starts after integration of viral DNA. The integration results in the inactivation of genes coding for the capsid proteins L1 and L2 and loss of E2 repressor function leads to deregulation of the E6/E7 open reading frame installing continuously overexpression of the two early proteins E6 and E7 that will lead to gradually loss of the normal cellular differentiation and the development of the 20 carcinoma. E6 and E7 overcome normal cell cycle by inactivating major tumor suppressor proteins, p53 and pRB, the retinoblastoma gene product, respectively.

Carcinoma of the cervix is common in women and develops through a pre-cancerous intermediate stage to the invasive carcinoma which frequently leads to death. The intermediate stages of the disease is known as cervical intraepithelial 25 neoplasia and is graded I to III in terms of increasing severity (*CIN I-III*).

Clinically, HPV infection of the female anogenital tract manifests as cervical flat condylomas, the hallmark of which is the koilocytosis affecting predominantly the superficial and intermediate cells of the cervical squamous epithelium.

Koilocytes which are the consequence of a cytopathic effect of the virus, 30 appear as multinucleated cells with a perinuclear clear halo. The epithelium is thickened with abnormal keratinisation responsible for the warty appearance of the lesion.

Such flat condylomas when positive for the HPV 16 or 18 serotypes, are high-risk factors for the evolution toward cervical intraepithelial neoplasia (CIN) and carcinoma in situ (CIS) which are themselves regarded as precursor lesions of invasive cervix carcinoma.

5        The natural history of oncogenic HPV infection presents three consecutive phases, namely:

- (1) a latent infection phase,
- (2) a phase of intranuclear viral replication with product of complete virions, which corresponds to the occurrence of koilocytes. At this stage, the HPV is producing its full range of proteins including E2, E5, E6, E7, L1 and L2.
- (3) a phase of viral integration into the cellular genome, which triggers the onset of malignant transformation, and corresponds to CIN II and CIN III/CIS with progressive disappearance of koilocytes. At this stage, the expression of E2 is down-regulated, the expression of E6 and E7 is enhanced. Between CIN II/III and CIN III / Cervix carcinoma the viral DNA changes from being episomal in the basal cells to integration of E6 and E7 genes only (tumoral cells). 85% of all cervix carcinomas are squamous cell carcinomas most predominantly related to the HPV16 serotype. 10% and 5% are adenocarcinomas and adenosquamous cell carcinomas respectively, and both types are predominantly related to HPV 18 serotype. Nevertheless other oncogenic HPV's exist.

20      International Patent Application No. WO 96/19496 discloses variants of human papilloma virus E6 and E7 proteins, particularly fusion proteins of E6/E7 with a deletion in both the E6 and E7 proteins. These deletion fusion proteins are said to be immunogenic.

25      Immunomodulatory oligonucleotides contain unmethylated CpG dinucleotides ("CpG") and are known (WO 96/02555, EP 468520). CpG is an abbreviation for cytosine-guanosine dinucleotide motifs present in DNA. Historically, it was observed that the DNA fraction of BCG could exert an anti-tumour effect. In further studies, synthetic oligonucleotides derived from BCG gene sequences were shown to be capable of inducing immunostimulatory effects (both in vitro and in vivo). The authors of these studies concluded that certain palindromic sequences, including a central CG motif, carried this activity. The central role of the CG motif in

immunostimulation was later elucidated in a publication by Krieg, Nature 374, p546 1995. Detailed analysis has shown that the CG motif has to be in a certain sequence context, and that such sequences are common in bacterial DNA but are rare in vertebrate DNA.

5 It is currently believed that this evolutionary difference allows the vertebrate immune system to detect the presence of bacterial DNA (as occurring during an infection) leading consequently to the stimulation of the immune system. The immunostimulatory sequence as defined by Krieg is:

Purine Purine CG pyrimidine pyrimidine and where the CG motif is not 10 methylated. In certain combinations of the six nucleotides a palindromic sequence is present. Several of these motifs, either as repeats of one motif or a combination of different motifs, can be present in the same oligonucleotide. The presence of one or more of these immunostimulatory sequence containing oligonucleotides can activate various immune subsets, including natural killer cells (which produce interferon  $\gamma$  and 15 have cytolytic activity) and macrophages (Wooldridge et al Vol 89 (no. 8), 1977). Although other unmethylated CpG containing sequences not having this consensus sequence have now been shown to be immunomodulatory.

The present invention provides compositions comprising either an E6 or/and 20 E7 or an E6/E7 fusion protein optionally linked to an immunological fusion partner having T cell epitopes, and adjuvanted with an immunomodulatory CpG containing oligonucleotide.

In a preferred form of the invention, the immunological fusion partner is derived from protein D of Haemophilus influenza B. Preferably the protein D derivative comprises approximately the first 1/3 of the protein, in particular 25 approximately the first N-terminal 100-110 amino acids. The protein D may be lipidated (Lipo Protein D). Other immunological fusion partners include the non-structural protein from influenzae virus, NS1 (hemagglutinin). Typically the N terminal 81 amino acids are utilised, although different fragments may be used provided they include T-helper epitopes.

30 In another embodiment the immunological fusion partner is the protein known as LYTA. Preferably the C terminal portion of the molecule is used. Lyta is derived from Streptococcus pneumoniae which synthesizes an N-acetyl-L-alanine amidase,

amidase LYTA, (coded by the *lytA* gen {Gene, 43 (1986) page 265-272} an autolysin that specifically degrades certain bonds in the peptidoglycan backbone. The C-terminal domain of the LYTA protein is responsible for the affinity to the choline or to some choline analogues such as DEAE. This property has been exploited for the development of *E.coli* C-LYTA expressing plasmids useful for expression of fusion proteins. Purification of hybrid proteins containing the C-LYTA fragment at its amino terminus has been described {Biotechnology: 10, (1992) page 795-798}. As used herein a preferred embodiment utilises the repeat portion of the Lyta molecule found in the C terminal end starting at residue 178. A particularly preferred form incorporates residues 188 - 305.

Accordingly, the present invention in preferred embodiment provides compositions comprising an immunomodulatory CpG oligonucleotide and a fusion proteins comprising Protein D - E6 from HPV 16, Protein D - E7 from HPV 16 Protein D - E7 from HPV 18, Protein D - E6 from HPV 18, and Protein D E6 E7 from both HPV 16 and 18. The protein D part preferably comprises the first 1/3 of protein D. It will be appreciated that other E6 and E7 proteins may be utilised from other HPV subtypes.

The proteins utilised in the present invention preferably are expressed in *E. coli*. In a preferred embodiment the proteins are expressed with a Histidine tail comprising between 5 to 9 and preferably six Histidine residues. These are advantageous in aiding purification.

The protein E7 may in a preferred embodiment carry one or several mutations in the binding site for the rb (retinoblastoma gene product) and hence eliminate any potential transforming capacity. Preferred mutations for HPV 16 E7 involve replacing Cys<sub>24</sub> with Glycine, or Glutamic acid<sub>26</sub> with Glutamine. In a preferred embodiment the E7 protein contains both these mutations.

Preferred mutations for the HPV 18 E<sub>7</sub> involve replacing Cys<sub>27</sub> with Glycine and/or Glutamic acid<sub>29</sub> with Glutamine. Again preferably both mutations are present.

Single or double mutations may also be introduced p53 region of E<sub>6</sub> to eliminate any potential transforming ability.

In a further embodiment of the invention there is provided and E6 E7 fusion protein from HPV linked to an immunological fusion partner and a CpG immunomodulatory oligonucleotide.

The vaccine of the present invention preferentially induce a TH1 immune response.

Two main types of Helper T cells have been characterized TH1 and TH2, which differ in the type of cytokines they secrete. These cytokines can be considered as the driving force behind the development of 2 different types of immune response : TH1-type of immune response is associated with cell mediated effector mechanisms such as production of the INF- $\gamma$  and IL-2 cytokines by T-lymphocytes. INF- $\gamma$  which in turn can activate other cells and induce them to secrete other important cytokines and mediators (INF- $\gamma$  - activated NK cells produce IL12, IL2-activated NK cells are transformed into lymphokine activated killer cell ( LAK), INF- $\gamma$ -activated macrophages secrete inflammatory mediators like TNFa, IL1, IL6 and release nitric oxyde , IL2 can provide help for the differentiation of antigen specific, haplotype restricted cytotoxic T lymphocytes (CTL). At the antibody level , in mice, Th1-type of immune response is also associated with the generation of antibodies of the IgG2 isotype (IgG2a in Balb/c mice and IgG2b in C57BL/6 mice) .

The Th2-type of immune response is associated with a humoral immune response to the antigen. with the production of cytokines like IL4, IL5, IL6, IL10 and by the generation of a broad range of immunoglobulin isotypes including in mice IgG1, IgA, and IgM.

In man the distinction of Th1 and Th2-type immune responses is not absolute. An individual will support an immune response which is predominantly Th1 or predominantly Th2. However, it is often convenient to consider the families of cytokines in terms of that described in murine CD4 +ve T cell clones by Mosmann and Coffman (*Mosmann, T.R. and Coffman, R.L. (1989) TH1 and TH2 cells: different patterns of lymphokine secretion lead to different functional properties. Annual Review of Immunology, 7, p145-173*).

In the human TH1 type of response is also associated with the presence of cytokine (IFNg and IL2) eventually with the presence of CT1 and IgG2 isotypes in mice correspond to IgG1 type antibodies

This type 1 phenotype is of particular importance in protecting against viral and intracellular bacterial infections as well as in the treatment of cancer.

To manufacture the proteins used in the invention by recombinant techniques, an expression strategy can be used which involves fusion of E7, E6 or E6/E7 fusion to the 1/3-N-terminal portion of protein D from *Haemophilus influenzae* B, an immunological fusion partner providing T cell helper epitopes. An affinity polyhistidine tail is engineered at the carboxy terminus of the fusion protein allowing for simplified purification. Such recombinant antigen is overexpressed in *E. coli* as insoluble protein.

The proteins of the invention may be coexpressed with thioredoxin in trans (TIT). Coexpression of thioredoxin in trans versus in cis is preferred to keep antigen free of thioredoxin without the need for protease. Thioredoxin coexpression eases the solubilisation of the proteins of the invention. Thioredoxin coexpression has also a significant impact on protein purification yield, on purified-protein solubility and quality.

The replicable expression vectors may be prepared in accordance with the invention, by cleaving a vector compatible with the host cell to provide a linear DNA segment having an intact replicon, and combining said linear segment with one or more DNA molecules which, together with said linear segment encode the desired product, such as the DNA polymer encoding the protein of the invention, or derivative thereof, under ligating conditions.

Thus, the DNA polymer may be preformed or formed during the construction of the vector, as desired.

The choice of vector will be determined in part by the host cell, which may be prokaryotic or eukaryotic but preferably is *E. coli*. Suitable vectors include plasmids, bacteriophages, cosmids and recombinant viruses.

The preparation of the replicable expression vector may be carried out conventionally with appropriate enzymes for restriction, polymerisation and ligation of the DNA, by procedures described in, for example, Maniatis *et al.* cited above.

The recombinant host cell is prepared, in accordance with the invention, by transforming a host cell with a replicable expression vector of the invention under transforming conditions. Suitable transforming conditions are conventional and are

described in, for example, Maniatis *et al.* cited above, or "DNA Cloning" Vol. II, D.M. Glover ed., IRL Press Ltd, 1985.

The choice of transforming conditions is determined by the host cell. Thus, a bacterial host such as *E. coli* may be treated with a solution of  $\text{CaCl}_2$  (Cohen *et al.*, 5 Proc. Nat. Acad. Sci., 1973, 69, 2110) or with a solution comprising a mixture of  $\text{RbCl}$ ,  $\text{MnCl}_2$ , potassium acetate and glycerol, and then with 3-[N-morpholino]-propane-sulphonic acid,  $\text{RbCl}$  and glycerol. Mammalian cells in culture may be transformed by calcium co-precipitation of the vector DNA onto the cells. The invention also extends to a host cell transformed with a replicable expression vector of 10 the invention.

Culturing the transformed host cell under conditions permitting expression of the DNA polymer is carried out conventionally, as described in, for example, Maniatis *et al.* and "DNA Cloning" cited above. Thus, preferably the cell is supplied with nutrient and cultured at a temperature below 50°C.

15 The product is recovered by conventional methods according to the host cell. Thus, where the host cell is bacterial, such as *E. coli* it may be lysed physically, chemically or enzymatically and the protein product isolated from the resulting lysate. Where the host cell is mammalian, the product may generally be isolated from the 20 nutrient medium or from cell free extracts. Conventional protein isolation techniques include selective precipitation, adsorption chromatography, and affinity chromatography including a monoclonal antibody affinity column.

When the proteins of the present invention are expressed with a histidine tail (His tag). The proteins can easily be purified by affinity chromatography using an ion metal affinity chromatography column (IMAC) column.

25 A second chromatographic step, such as Q-sepharose may be utilised either before or after the IMAC column to yield highly purified protein. If the immunological fusion partner is C-LYTA, then it is possible to exploit the affinity of CLYTA for choline and/or DEAE to purify this product. Products containing both C-LYTA and his tags can be easily and efficiently purified in a two step process 30 involving differential affinity chromatography. One step involves the affinity of the His tag to IMAC columns, the other involves the affinity of the C-terminal domain of LYTA for choline or DEAE.

A preferred vaccine composition comprises at least Protein D - E6 from HPV 16 or derivative thereof together with Protein D - E7 from HPV 16. Alternatively the E6 and E7 may be presented in a single molecule, preferably a Protein D E6/E7 fusion. Such vaccine may optionally contain either or both E6 and E7 proteins from HPV 18, preferably in the form of a Protein D - E6 or Protein D - E7 fusion protein or Protein D E6/E7 fusion protein. The vaccines of the present invention may contain other HPV antigens from HPV 16 or 18. In particular, the vaccine may contain L1 or L2 antigen monomers. Alternatively such L1 or L2 antigens may be presented together as a virus like particle or the L1 alone protein may be presented as virus like particle or capsomer structure. Such antigens, virus like particles and capsomer are per se known. See for example WO94/00152, WO94/20137, WO94/05792, and WO93/02184. Additional early proteins may be included such as E2 or preferably E5 for example. The vaccine of the present invention may additionally comprise antigens from other HPV strains, preferably from strains HPV 6, 11, 31 or 33.

Vaccine preparation is generally described in Vaccine Design - The subunit and adjuvant approach (Ed. Powell and Newman) Pharmaceutical Biotechnology Vol. 6 Plenum Press 1995. Encapsulation within liposomes is described by Fullerton, US Patent 4,235,877.

The preferred oligonucleotides preferably contain two or more CpG motifs separated by six or more nucleotides. The oligonucleotides of the present invention are typically deoxynucleotides. In a preferred embodiment the internucleotide in the oligonucleotide is phosphorodithioate, or more preferably a phosphorothioate bond, although phosphodiester and other internucleotide bonds are within the scope of the invention including oligonucleotides with mixed internucleotide linkages.

Preferred oligonucleotides have the following sequences: The sequences preferably contain all phosphorothioate modified internucleotide linkages.

OLIGO 1: TCC ATG ACG TTC CTG ACG TT

OLIGO 2: TCT CCC AGC GTG CGC CAT

OLIGO 3: ACC GAT GAC GTC GCC GGT GAC GGC ACC ACG

The CpG oligonucleotides utilised in the present invention may be synthesized by any method known in the art (eg EP 468520). Conveniently, such oligonucleotides may be synthesized utilising an automated synthesizer. Methods for producing

phosphorothioate oligonucleotides or phosphorodithioate are described in US5,666,153, US5,278,302 and WO95/26204.

The invention will be further described by reference to the following examples:

5   **EXAMPLE I: Construction of an E. coli strain expressing fusion Protein-D1/3 - E7 -His (HPV16)**

1) - **Construction of expression plasmid**

a) - Plasmid pMG MCS prot D1/3 (= pRIT14589) is a derivative of pMG81 (described in UK patent application n° 951 3261.9 published as WO97/01640) in which the codons 4-81 of NS1 coding region from Influenza were replaced by the codons corresponding to residues Ser 20 → Thr127 of mature protein D of Haemophilus Influenzae strain 772, biotype 2 (H. Janson *et al.*, 1991, Infection and Immunity, Jan. p.119-125). The sequence of Prot-D1/3 is followed by a multiple cloning site (11 residues) and a coding region for a C-terminal histidine tail (6 His).

15   This plasmid is used to express the fusion protein D1/3-E7-His.

b) - HPV genomic E6 and E7 sequences type HPV 16 (See Dorf *et al.*, Virology 1985, 145, p. 181-185) were amplified from HPV 16 full length genome cloned in pBR322 (obtained from Deutsches Krebsforschungszentrum (DKFZ), Referenzzentrum für human pathogen Papillomaviruses - D 69120 - Heidelberg) and 20   were subcloned into pUC19 to give TCA 301 (= pRIT14462).

**Construction of plasmid TCA 308 (= pRIT14501): a plasmid expressing the fusion Protein-D1/3-E7-His**

The nucleotides sequences corresponding to amino acids 1 → 98 of E7 protein are amplified from pRIT14462. During the polymerase chain reaction, Ncol and SpeI restriction sites were generated at the 5' and 3' ends of the E7 sequences 25   allowing insertion into the same sites of plasmid pMGMCS Prot D1/3 to give plasmid TCA308 (= pRIT14501). The insert was sequenced to verify that no modification had been generated during the polymerase chain reaction. The sequence for the fusion protein-D1/3-E7-His (HPV 16) is described in sequence ID No.1 and the coding 30   sequence in ID No.2.

**2) - Transformation of AR58 strain**

Plasmid pRIT14501 was introduced into E. coli AR58 (Mott *et al.*, 1985, Proc. Natl. Acad. Sci., 82:88) a defective λ lysogen containing a thermosensitive repressor of the λ pL promoter.

**3) - Growth and induction of bacterial strain - Expression of Prot-D1/3-E7-His**

Cells of AR58 transformed with plasmid pRIT14501 were grown in 100 ml of LB medium supplemented with 50 µgr/ml of Kanamycin at 30°C. During the logarithmic phase of growth bacteria were shifted to 39°C to inactivate the λ repressor and turn on the synthesis of protein D1/3-E7-His. The incubation at 39°C was continued for 4 hours. Bacteria were pelleted and stored at -20°C.

**10 EXAMPLE II: Construction of an *E.coli* strain expressing fusion Protein-D1/3-E6-his / HPV16**

**1. Construction of expression plasmid**

a) Plasmid pMG MCS prot D1/3 (= pRIT14589) is a derivative of pMG81 (described in WO97/01640 in which the codons 4-81 of NS1 coding region from Influenza were replaced by the codons corresponding to residues Ser 20 → Thr 127 of mature protein D of Haemophilus Influenzae strain 772, biotype 2 (H. Janson *et al.*, 1991, Infection and Immunity, Jan. p.119-125). The sequence of Prot-D1/3 is followed by a multiple cloning site (11 residues) and a coding region for a C-terminal histidine tail (6 His). This plasmid is used to express the fusion protein D1/3-E6-his.

b) HPV genomic E6 and E7 sequences type HPV16 (Seedorf *et al.*, Virology 1985, 145, p.181-185) were amplified from HPV16 full length genome cloned in pBR322 (obtained from Deutsches Krebsforschungszentrum (DKFZ), Referenzzentrum für human pathogen Papillomaviruses -

c) D 69120 - Heidelberg) and were subcloned into pUC19 to give TCA 301 (= pRIT14462).

**25 Construction of plasmid TCA 307 (=pRIT14497) : a plasmid expressing the fusion Protein-D1/3-E6-His /HPV16**

The nucleotides sequences corresponding to amino acid.

1 → 151 of E6 protein were amplified from pRIT14462. During the polymerase chain reaction, NcoI and SpeI restriction sites were generated at the 5' and 3' ends of the E6 sequences allowing insertion into the same sites of plasmid pMGMCS Prot D1/3 to give plasmid TCA307 (= pRIT14497). The insert was

sequenced to verify that no modification had been generated during the polymerase chain reaction. The protein and coding sequence for the fusion protein-D1/3-E6-His is described in sequence ID No.3 and 4.

## 2. Transformation of AR58 strain

5 Plasmid pRIT14497 was introduced into *E. coli* AR58 (Mott et al., 1985, Proc. Natl. Acad. Sci., 82:88) a defective  $\lambda$  lysogen containing a thermosensitive repressor of the  $\lambda$  pL promoter.

## 3. Growth and induction of bacterial strain - Expression of Prot-D1/3-E6-His

10 Cells of AR58 transformed with plasmid pRIT14497 were grown in 100 ml of LB medium supplemented with 50  $\mu$ gr/ml of Kanamycin at 30°C. During the logarithmic phase of growth bacteria were shifted to 39°C to inactivate the  $\lambda$  repressor and turn on the synthesis of protein D1/3-E6-his. The incubation at 39°C was continued for 4 hours. Bacteria were pelleted and stored at -20C.

## 4. Characterization of fusion Protein D1/3-E6-his (HPV 16)

### 15 Preparation of extracts

Frozen cells are thawed and resuspended in 10 ml of PBS buffer. Cells are broken in a French pressure cell press SLM Aminco at 20.000 psi (three passages). The extract is centrifuged at 16.000 g for 30 minutes at 4°C.

### Analysis on Coomassie-stained SDS-polyacrylamide gels and Western blots

20 After centrifugation of extracts described above, aliquots of supernatant and pellet were analysed by SDS-polyacrylamide gel electrophoresis and Western blotting.

25 A major band of about 32 kDa, localized in the pellet fraction, was visualised by Coomassie stained gels and identified in Western blots by rabbit polyclonal anti-protein-D and by Ni-NTA conjugate coupled to calf intestinal alkaline phosphatase (Qiagen cat. n° 34510) which detects accessible histidine tail. The level of expression represents about 5 % of total protein.

## 5. Coexpression with thioredoxin

In an analagons fashion to the expression of prot D 1/3 E7 His from HPV 18  
30 (example IX) an *E.coli* strain AR58 was transformed with a plasmid encoding thioredoxin and protein D 1/3 E7 His (HPV 16).

**EXAMPLE III: Construction of an *E. coli* strain expressing fusion Protein-D1/3-E6E7-his / HPV16**

**1. Construction of expression plasmid**

- a) Plasmid pMG MCS prot D1/3 (= pRIT14589) is a derivative of pMG81 (described Supra) in which the codons 4-81 of NS1 coding region from Influenza were replaced by the codons corresponding to residues Ser 20 → Thr 127 of mature protein D of Haemophilus Influenzae strain 772, biotype 2 (H. Janson *et al.*, 1991, Infection and Immunity, Jan. p.119-125). The sequence of Prot-D1/3 is followed by a multiple cloning site (11 residues) and a coding region for a C-terminal histidine tail (6 His).
- This plasmid is used to express the fusion protein D1/3-E6E7-his.
- b) HPV genomic E6 and E7 sequences type HPV16 (Seedorf *et al.*, Virology 1985, 145, p.181-185) were amplified from HPV16 full length genome cloned in pBR322 (obtained from Deutsches Krebsforschungszentrum (DKFZ), Referenzzentrum für human pathogen Papillomaviruses - D 69120 - Heidelberg) and were subcloned into pUC19 to give TCA 301 (= pRIT14462).
- c) The coding sequences for E6 and E7 in TCA301 (= pRIT 14462) were modified with a synthetic oligonucleotides adaptor (inserted between Afl III and Nsi I sites) introducing a deletion of 5 nucleotides between E6 and E7 genes to remove the stop codon of E6 and create fused E6 and E7 coding sequences in the plasmid TCA309( = pRIT 14556 ).

**Construction of plasmid TCA 311(= pRIT14512) : a plasmid expressing the fusion Protein-D1/3-E6E7-His /HPV16**

The nucleotides sequences corresponding to amino acids 1 → 249 of fused E6E7 protein were amplified from pRIT14556. During the polymerase chain reaction, NcoI and SpeI restriction sites were generated at the 5' and 3' ends of the E6E7 fused sequences allowing insertion into the same sites of plasmid pMGMCS Prot D1/3 to give plasmid TCA311 (= pRIT14512). The insert was sequenced to verify that no modification had been generated during the polymerase chain reaction. The protein and coding sequence for the fusion protein-D E6/E7 1/3-His is described sequence ID No. 5 and 6.

**2. Transformation of AR58 strain**

Plasmid pRIT14512 was introduced into *E. coli* AR58 (Mott et al., 1985, Proc. Natl. Acad. Sci., 82:88) a defective  $\lambda$  lysogen containing a thermosensitive repressor of the  $\lambda$  pL promoter.

**3. Growth and induction of bacterial strain - Expression of Prot-D1/3-E6E7-His**

Cells of AR58 transformed with plasmid pRIT14512 were grown in 100 ml of LB medium supplemented with 50  $\mu$ gr/ml of Kanamycin at 30°C. During the logarithmic phase of growth bacteria were shifted to 39°C to inactivate the  $\lambda$  repressor and turn on the synthesis of protein D1/3-E6E7-his. The incubation at 39°C was continued for 4 hours. Bacteria were pelleted and stored at -20C.

**10 4. Characterization of fusion Protein D1/3-E6E7-his**

Frozen cells are thawed and resuspended in 10 ml of PBS buffer. Cells are broken in a French pressure cell press SLM Aminco at 20.000 psi (three passages). The extract is centrifuged at 16.000 g for 30 minutes at 4°C.

After centrifugation of extracts described above, aliquots of supernatant and pellet were analysed by SDS-polyacrylamide gel electrophoresis and Western blotting.

A major band of about 48 kDa, localized in the pellet fraction, was visualised by Coomassie stained gels and identified in Western blots by rabbit polyclonal anti-protein-D and by Ni-NTA conjugate coupled to calf intestinal alkaline phosphatase (Qiagen cat. n° 34510) which detects accessible histidine tail. The level of expression represents about 1 % of total protein.

**EXAMPLE: IV**

In an analogous fashion the fusion protein of Lipo D 1/3 and E6-E7 from HPV16 was expressed in *E.coli* in the presence of thioredoxin.

The N-terminal of the pre-protein (388 aa) contains MDP residues followed by 16 amino acids of signal peptide of lipoprotein D (from Haemophilus Influenzae) which is cleaved in vivo to give the mature protein (370 aa). Lipoprotein portion (aa 1 to 127) is followed by the proteins E6 and E7 in fusion. The C terminal of the protein is elongated by TSGHHHHHH.

**30 EXAMPLE V: Construction of *E.coli* strain B1002 expressing fusion ProtD1/3-E7**

**Mutated (cys24->gly,glu26->gln ) type HPV16**

**1)-Construction of expression plasmid****Starting material:**

a) - Plasmid pRIT 14501 (= TCA 308) which codes for fusion ProtD1/3-E7 -His

b) - Plasmid LITMUS 28 ( New England Biolabs cat n° 306-28 ), a cloning vector

5 pUC-derived

c) - Plasmid pMG MCS ProtD1/3 ( pRIT 14589 ) , a derivative of pMG81 (described Supra) in which the codons 4-81 of NS1 coding region from Influenza were replaced by the codons corresponding to residues Ser 20 → Thr 127 of mature protein D of Haemophilus Influenzae strain 772, biotype 2 (H. Janson *et al.*, 1991, Infection and

10 Immunity, Jan. p.119-125). The sequence of Prot-D1/3 is followed by a multiple cloning site (11 residues) and a coding region for a C-terminal histidine tail (6 His)

**Construction of plasmid pRIT 14733( =TCA347): a plasmid expressing the fusion Protein-D1/3-E7 mutated ( cys24->gly ,glu26->gln ) with His tail**

The NcoI - XbaI fragment from pRIT 14501 (=TCA 308 ), bearing the coding sequence of E7 gene from HPV16 , elongated with an His tail , was subcloned in an intermediate vector Litmus 28 useful for mutagenesis to give pRIT 14909 (=TCA337) Double mutations **cys24-->gly** ( Edmonds and Vousden , J.Virology 63 : 2650 (1989) and **glu26-->gln** ( Phelps *et al* , J.Virology 66: 2418-27 ( 1992 ) were chosen to impair the binding to the antioncogene product of Retinoblastome gene ( pRB ).

15 The introduction of mutations in E7 gene was realized with the kit " Quick Change Site directed Mutagenesis ( Stratagene cat n° 200518) to give plasmid pRIT 14681(=TCA343) .After verification of presence of mutations and integrity of the complete E7 gene by sequencing , the mutated E7 gene was introduced into vector pRIT 14589 ( = pMG MCS ProtD1/3 ) to give plasmid pRIT 14733 (=TCA347) 20 protein and coding sequence.

25 The sequence for the fusion protein-D1/3-E 7 mutated ( cys24->gly, glu26->gln ) -His is described in sequence ID No. 7 and 8.

**2)-Construction of strain B1002 expressing ProtD1/3-E7mutated (cys 24-->gly , glu26-->gln )-His /HPV16**

30 Plasmid pRIT 14733 was introduced into *E.coli AR58* ( Mott *et al.* ,1985,

Proc. Natl. Acad. Sci. , 82:88 ) a defective  $\lambda$  lysogen containing a thermosensitive repressor of the  $\lambda$  pL promoter ,to give strain B1002 , by selection for transformants resistant to kanamycin

3)-Growth and induction of bacterial strain B1002 - Expression of ProtD1/3-E7 mutated (cys 24->gly , glu26->gln )-His /HPV16

Cells of AR58 transformed with plasmid pRIT 14733 ( B1002 strain ) were grown at 30°C in 100 ml of LB medium supplemented with 50  $\mu$ gr /ml of Kanamycin. During the logarithmic phase of growth bacteria were shifted to 39°C to inactivate the  $\lambda$  repressor and turn on the synthesis of ProtD1/3-E7 mutated -His /HPV16 . The 10 incubation at 39°C was continued for 4 hours . Bacteria were pelleted and stored at - 20°C.

4)-Characterization of fusion ProtD1/3-E7 mut (cys24->gly, glu26->gln)- His type HPV16.

Frozen cells were thawed and resuspended in 10 ml of PBS buffer.Cells were 15 broken in a French Pressure cell press SLM Aminco at 20 000 psi ( three passages) . The extract was centrifuged at 16000 g for 30 minutes at 4°C.

After centrifugation of extracts described above, aliquots of supernatant and pellet were analysed by SDS-polyacrylamide gel electrophoresis and Western blotting.

A major band of about 33 kDa, localized in the pellet fraction, was visualised 20 by Coomassie stained gels and identified in Western blots by rabbit polyclonal 22 J 70 anti-protein D, by monoclonal anti E7 /HPV16 from Zymed and by Ni-NTA conjugate coupled to calf intestinal alkaline phosphatase (Qiagen cat. n° 34510) which detects accessible histidine tail. The level of expression represents about 3 to5 % of 25 total protein.

Cells of B1002 were separated from the culture broth by centrifugation. The concentrated cells of B1002 were stored at -65°C.

**EXAMPLE VI: Construction of an *E. coli* strain expressing fusion clyta-E6-his (HPV 16)**

30 **1. Construction of expression plasmid**

a) -Plasmid pRIT14497 (= TCA307), that codes for fusion ProtD1/3-E6-His /HPV16

b)-Plasmid pRIT14661 (= DVA2), an intermediate vector containing the coding sequence for the 117 C-terminal codons of LytA of *Streptococcus Pneumoniae*. Lyta is derived from *Streptococcus pneumoniae* which synthesize an N-acetyl-L-alanine amidase, amidase LYTA. (coded by the *lytA* gene {Gene, 43 (1986) pag 265-272} an autolysin that specifically degrades certain bonds in the peptidoglycan backbone .  
5 The C-terminal domain of the LYTA protein is responsible for the affinity to the choline or to some choline analogues such as DEAE.

**1.b Construction of plasmid pRIT14634 (=TCA332): a plasmid expressing the fusion clyta-E6-His /HPV16**

- 10 a)The first step was the purification of the large NcoI-AfII restriction fragment from plasmid pRIT14497 and the purification of the small AfIII-AfIII restriction fragment from pRIT14661  
b)The second step was linking of clyta sequences to the E7-His sequences (NcoI and AfIII are compatible restriction sites) that gave rise to the plasmid pRIT 14634  
15 (=TCA332), coding for the fusion protein clyta-E6-His under the control of the pL promoter.

The protein and coding sequence for the fusion protein clyta-E6-His is described sequence ID No. 9 and 10.

**Transformation of AR58 strain**

- 20 Plasmid pRIT14634 was introduced into *E. coli* AR58 (Mott et al., 1985, Proc. Natl. Acad. Sci., 82:88) a defective λ lysogen containing a thermosensitive repressor of the λ pL promoter.

**Growth and induction of bacterial strain - Expression of clyta-E6-His**

- 25 Cells of AR58 transformed with plasmid pRIT14634 were grown in 100 ml of LB medium supplemented with 50 µgr/ml of Kanamycin at 30°C. During the logarithmic phase of growth bacteria were shifted to 39°C to inactivate the λ repressor and turn on the synthesis of protein clyta-E6-his. The incubation at 39°C was continued for 4 hours. Bacteria were pelleted and stored at -20°C.

**4. Characterization of fusion clyta-E6-his**

- 30 Frozen cells were thawed and resuspended in 10 ml of PBS buffer. Cells were broken in a French pressure cell press SLM Aminco at 20.000 psi (three passages). The extract was centrifuged at 16.000 g for 30 minutes at 4°C. After centrifugation of

extracts described above, aliquots of supernatant and pellet were analysed by SDS-polyacrylamide gel electrophoresis and Western blotting.

A major band of about 33 kDa, localized in the pellet fraction, was visualised by Coomassie stained gels and identified in Western blots by rabbit polyclonal anti-clyta antibodies and by Ni-NTA conjugate coupled to calf intestinal alkaline phosphatase (Qiagen cat. n° 34510) which detects accessible histidine tail. The level of expression represents about 3 % of total protein.

**EXAMPLE VII: Construction of an *E. coli* strain expressing fusion clyta-E7-his (HPV 16)**

10    **1. Construction of expression plasmid**

1.a Starting materials

- a) -Plasmid pRIT14501 (= TCA308), that codes for fusion ProtD1/3-E7-His /HPV16
- b)-Plasmid pRIT14661 (= DVA2), an intermediate vector containing the coding sequence for the 117 C-terminal codons of LytA of Streptococcus Pneumoniae.

15    **1.b Construction of plasmid pRIT14626 (=TCA330): a plasmid expressing the fusion clyta-E7-His / HPV16**

- a) The first step was the purification of the large Ncol-AflII restriction fragment from plasmid pRIT14501 and the purification of the small AflIII-AflIII restriction fragment from pRIT14661
- 20    b) The second step was linking of clyta sequences to the E7-His sequences (Ncol and AflIII are compatible restriction sites) that gave rise to the plasmid pRIT 14626 (=TCA330), coding for the fusion protein clyta-E7-His under the control of the pL promoter.

The protein and coding sequence for the fusion protein clyta-E7-His is  
25    described in sequence ID No. 11 and 12.

**2. Transformation of AR58 strain**

Plasmid pRIT14626 was introduced into *E. coli* AR58 (Mott et al., 1985, Proc. Natl. Acad. Sci., 82:88) a defective λ lysogen containing a thermosensitive repressor of the λ pL promoter.

30    **3. Growth and induction of bacterial strain - Expression of clyta-E7-His**

Cells of AR58 transformed with plasmid pRIT14626 were grown in 100 ml of LB medium supplemented with 50 µgr/ml of Kanamycin at 30°C. During the

logarithmic phase of growth bacteria were shifted to 39°C to inactivate the λ repressor and turn on the synthesis of protein clyta-E7-his. The incubation at 39°C was continued for 4 hours. Bacteria were pelleted and stored at -20°C.

#### 4. Characterization of fusion clyta-E7-his

- 5       Frozen cells were thawed and resuspended in 10 ml of PBS buffer. Cells were broken in a French pressure cell press SLM Aminco at 20.000 psi (three passages). The extract was centrifuged at 16.000 g for 30 minutes at 4°C. After centrifugation of extracts described above, aliquots of supernatant and pellet were analysed by SDS-polyacrylamide gel electrophoresis and Western blotting.
- 10      A major band of about 35 kDa, localized in the pellet fraction, was visualised by Coomassie stained gels and identified in Western blots by rabbit polyclonal anti-clyta antibodies and by Ni-NTA conjugate coupled to calf intestinal alkaline phosphatase (Qiagen cat. n° 34510) which detects accessible histidine tail. The level of expression represents about 5 % of total protein.

15      **EXAMPLE VIII: Construction of an *E. coli* strain expressing fusion clyta-E6E7-his (HPV 16)**

##### 1. Construction of expression plasmid

###### 1.a Starting materials

- a) -Plasmid pRIT14512 (= TCA311), that codes for fusion ProtD1/3-E6E7-His /HPV16
- b)-Plasmid pRIT14661 (= DVA2), an intermediate vector containing the coding sequence for the 117 C-terminal codons of LytA of Streptococcus Pneumoniae.

###### 1.b Construction of plasmid pRIT14629 (=TCA331): a plasmid expressing the fusion clyta-E6E7-His /HPV16

- 25      a)The first step was the purification of the large NcoI-AflII restriction fragment from plasmid pRIT14512 and the purification of the small AflII-AflIII restriction fragment from pRIT14661
- b)The second step was linking of clyta sequences to the E7-His sequences (NcoI and AflIII are compatible restriction sites)that gave rise to the plasmid pRIT 14629
- 30      (=TCA331), coding for the fusion protein clyta-E6E7-His under the control of the pL promoter.

The protein and coding sequence for the fusion protein clyta-E6E7-His is sequenced ID No. 13 and 14.

## 2. Transformation of AR58 strain

Plasmid pRIT14629 was introduced into *E. coli* AR58 (Mott et al., 1985, Proc. Natl. Acad. Sci., 82:88) a defective λ lysogen containing a thermosensitive repressor of the λ pL promoter.

## 3. Growth and induction of bacterial strain - Expression of clyta-E6E7-His

Cells of AR58 transformed with plasmid pRIT14629 were grown in 100 ml of LB medium supplemented with 50 µgr/ml of Kanamycin at 30°C. During the logarithmic phase of growth bacteria were shifted to 39°C to inactivate the λ repressor and turn on the synthesis of protein clyta-E6E7-his. The incubation at 39°C was continued for 4 hours. Bacteria were pelleted and stored at -20°C.

## 4. Characterization of fusion clyta-E6E7-his

Frozen cells were thawed and resuspended in 10 ml of PBS buffer. Cells were broken in a French pressure cell press SLM Aminco at 20.000 psi (three passages). The extract was centrifuged at 16.000 g for 30 minutes at 4°C.

After centrifugation of extracts described above, aliquots of supernatant and pellet were analysed by SDS-polyacrylamide gel electrophoresis and Western blotting.

A major band of about 48 kDa, localized in the pellet fraction, was visualised by Coomassie stained gels and identified in Western blots by rabbit polyclonal anti-clyta antibodies and by Ni-NTA conjugate coupled to calf intestinal alkaline phosphatase (Qiagen cat. n° 34510) which detects accessible histidine tail. The level of expression represents about 1 % of total protein.

## 25 EXAMPLE IX: Prot D1/3 E7 his (HPV 18) (E.Coli B1011)

Protein D1/3 E7 his HPV expressed with Thioredoxin inTrans (E.Coli B1012)

### 1) - Construction of expression plasmids

#### 1).a.Construction of plasmid TCA316(=pRIT 14532) a plasmid expressing the fusion Protein-D1/3-E7-His /HPV18

##### 30 Starting materials

a) - Plasmid pMG MCS prot D1/3 (= pRIT14589) is a derivative of pMG81 (described in UK patent application n° 951 3261.9 published as WO97/01640 in

which the codons 4-81 of NS1 coding region from Influenza were replaced by the codons corresponding to residues Ser 20 → Thr 127 of mature protein D of Haemophilus Influenzae strain 772, biotype 2 (H. Janson *et al.*, 1991, Infection and Immunity, Jan. p.119-125). The sequence of Prot-D1/3 is followed by a multiple cloning site (11 residues) and a coding region for a C-terminal histidine tail (6 His).

5 This plasmid is used to express the fusion protein D1/3-E7-his.

b) - HPV genomic E6 and E7 sequences of prototype HPV18(Cole et al.J.Mol.Biol.(1987)193,599-608) were amplified from HPV16 full length genome cloned in pBR322 (obtained from Deutsche Krebsforschungszentrum (DKFZ), Referenzzentrum für human pathogen Papillomaviruses - D 69120 - Heidelberg) and were subcloned into pUC19 to give TCA 302 (= pRIT14467).

10

#### **Construction of plasmid TCA 316(= pRIT14532)**

The nucleotides sequences corresponding to amino acids 1 → 105 of E7 protein were amplified from pRIT14467. During the polymerase chain reaction, NcoI and SpeI restriction sites were generated at the 5' and 3' ends of the E7 sequences allowing insertion into the same sites of plasmid pMGMCS Prot D1/3 to give plasmid TCA316 (= pRIT14532). The insert was sequenced and a modification versus E7/HPV18 prototype sequence was identified in E7 gene (nucleotide 128 G->A) generating a substitution of a glycine by a glutamic acid (aa 43 in E7 , position 156 in fusion protein). The protein and coding sequence for the fusion protein-D1/3-E7-His /HPV18 is set forth in sequence ID No. 15 and ID No. 16.

15

#### **1).b. Construction of plasmid TCA313 (=pRIT14523): a plasmid expressing thioredoxin**

##### **Starting materials**

20 a) - Plasmid pBBR1MCS4(Antoine R. and C.Locht,Mol.Microbiol. 1992,6,1785-1799 ; M.E.Kovach et al. Biotechniques 16, (5), 800-802 )which is compatible with plasmids containing ColE1 or P15a origins of replication.

b) - Plasmid pMG42 (described in WO93/04175) containing the sequence of promoter pL of Lambda phage

25 c) - Plasmid pTRX ( Invitrogen, kit Thiofusion K350-01) bearing the coding sequence for thioredoxin followed by AspA transcription terminator.

#### **Construction of plasmid TCA313(=pRIT14523)**

The fragment EcoRI-NdeI fragment from pMG42, bearing pL promoter and the NdeI-HindIII fragment from pTRX, bearing the coding sequence for thioredoxin followed by AspA terminator, were purified and ligated into the EcoRI and HindIII sites of plasmid vector pBBR1MCS4 to give plasmid TCA313(= pRIT14523).

5 The coding sequence for thioredoxin is described in ID No. 17.

## 2) - Transformation of AR58 strain

### 2).a. To obtain strain B1011 expressing ProtD1/3-E7-His/HPV18

Plasmid pRIT14532 was introduced into *E. coli* AR58 (Mott et al., 1985, Proc. Natl. Acad. Sci., 82:88) a defective  $\lambda$  lysogen containing a thermosensitive repressor of the 10  $\lambda$  pL promoter , by selection for transformants resistant to kanamycin.

### 2).b. Construction of strain B1012 expressing ProtD1/3-E7-His/HPV18 and thioredoxin

15 Plasmid pRIT14532 and pRIT14523 were introduced into *E. coli* AR58 (Mott et al., 1985, Proc. Natl. Acad. Sci., 82:88) a defective  $\lambda$  lysogen containing a thermosensitive repressor of the  $\lambda$  pL promoter ,by double selection for transformants 20 resistant to kanamycin and ampicillin.

## 3) - Growth and induction of bacterial strains B1011 and B1012 - Expression of Prot-D1/3-E7-His/HPV18 without and with thioredoxin in trans

Cells of AR58 transformed with plasmids pRIT14532 ( B1011 strain) and 20 Cells of AR58 transformed with plasmids pRIT14532 and pRIT14523 (B1012 strain) were grown at 30°C in 100 ml of LB medium supplemented with 50  $\mu$ gr/ml of Kanamycin for B1011 strain and supplemented 50  $\mu$ gr/ml of Kanamycin and 100  $\mu$ gr/ml of Ampicillin for B1012 strain . During the logarithmic phase of growth 25 bacteria were shifted to 39°C to inactivate the  $\lambda$  repressor and turn on the synthesis of protein D1/3-E7-his/HPV18 and thioredoxin. The incubation at 39°C was continued for 4 hours.

## Characterization of fusion Protein D1/3-E7-his /HPV18

### Preparation of extracts

Frozen cells are thawed and resuspended in 10 ml of PBS buffer. Cells are 30 broken in a French pressure cell press SLM Aminco at 20.000 psi (three passages). The extract is centrifuged at 16.000 g for 30 minutes at 4°C.

### Analysis on Coomassie-stained SDS-polyacrylamide gels and Western blots

After centrifugation of extracts described above, aliquots of supernatant and pellet were analysed by SDS-polyacrylamide gel electrophoresis and Western blotting.

The fusion protD1/3-E7-His (about 31 kDa) was visualised by Coomassie stained gels in the pellet fraction for strain B1011 and partially localized (30%) in the supernatant fraction for strain B1012 and was identified in Western blots by rabbit polyclonal anti-protein-D and by Ni-NTA conjugate coupled to calf intestinal alkaline phosphatase (Qiagen cat. n° 34510) which detects accessible histidine tail. The level of expression represents about 1-3% of total protein as shown on a Coomassie-stained SDS-polyacrylamide gel.

For the extract of strain B1012 the thioredoxin ( about 12 KDa) was visualised by coomassie stained gel in the supernatant and identified in western blots by monoclonal anti thioredoxin ( Invitrogen R920-25)

**EXAMPLE X: Construction of E.coli strain B1098 expressing fusion ProtD1/3-**

15      **E7**

**Mutated (cys27->gly,glu29->gln ) type HPV18**

**1)-Construction of expression plasmid**

**Starting material:**

- a) - Plasmid pRIT 14532 (= TCA 316) which codes for fusion ProtD1/3-E7 -His
- 20 b) - Plasmid LITMUS 28 ( New England Biolabs cat n° 306-28 ), a cloning vector pUC-derived
- c) - Plasmid pMG MCS ProtD1/3 ( pRIT 14589 ), a derivative of pMG81 (described supra) in which the codons 4-81 of NS1 coding region from Influenza were replaced by the codons corresponding to residues Ser 20 → Thr 127 of mature protein D of
- 25 Haemophilus Influenzae strain 772, biotype 2 (H. Janson *et al.*, 1991, Infection and Immunity, Jan. p.119-125). The sequence of Prot-D1/3 is followed by a multiple cloning site (11 residues) and a coding region for a C-terminal histidine tail (6 His)

**Construction of plasmid pRIT 14831( =TCA355): a plasmid expressing the fusion Protein-D1/3-E7 mutated ( cys27->gly ,glu29->gln ) with His tail**

30      The Ncol - XbaI fragment from pRIT 14532 (=TCA 316 ), bearing the coding sequence of E7 gene from HPV18 , elongated with an His tail , was subcloned in an intermediate vector Litmus 28 useful for mutagenesis to give pRIT 14910 (=TCA348)

By analogy with E7/HPV16 mutagenesis, double mutations **cys27-->gly** and **glu29-->gln** were chosen to impair the binding to the antioncogene product of Retinoblastome gene ( pRB ).

The introduction of mutations in E7 gene was realized with the kit " Quick Change Site directed Mutagenesis ( Stratagene cat n° 200518) . As the sequencing of pRIT14532 had pointed out the presence of a glutamic acid in position 43 of E7 instead of a glycine in the prototype sequence of HPV18 , a second cycle of mutagenesis was realized to introduce a glycine in position 43 . We obtained plasmid pRIT 14829 ( = TCA353 ) . After verification of presence of mutations and integrity of the complete E7 gene by sequencing , the mutated E7 gene was introduced into vector pRIT 14589 ( = pMG MCS ProtD1/3 ) to give plasmid pRIT 14831 ( =TCA355 ).

The protein and coding sequence for the fusion protein-D1/3-E 7 mutated (cys27->gly, glu29->gln ) -His is described in sequence ID No. 18 and 19.

**15 2)Construction of strain B1098 expressing ProtD1/3-E7mutated (cys 27-->gly , glu29-->gln )-His /HPV18**

Plasmid pRIT 14831 was introduced into *E.coli AR58* ( Mott et al. ,1985, Proc. Natl. Acad. Sci. , 82:88 ) a defective λ lysogen containing a thermosensitive repressor of the λ pL promoter ,to give strain B1098 , by selection for transformants 20 resistant to kanamycin.

**3)-Growth and induction of bacterial strain B1098 - Expression of ProtD1/3-E7 mutated (cys 27->gly , glu29->gln )-His /HPV18**

Cells of AR58 transformed with plasmid pRIT 14831 ( B1098 strain ) were grown at 30°C in 100 ml of LB medium supplemented with 50 µgr /ml of Kanamycin. 25 During the logarithmic phase of growth bacteria were shifted to 39°C to inactivate the λ repressor and turn on the synthesis of ProtD1/3-E7 mutated -His /HPV18 . The incubation at 39°C was continued for 4 hours. Bacteria were pelleted and stored at - 20°C.

**4)-Characterization of fusion ProtD1/3-E7 mut (cys24->gly, glu26->gln)- His type 30 HPV16**

Frozen cells were thawed and resuspended in 10 ml of PBS buffer. Cells were broken in a French Pressure cell press SLM Aminco at 20 000 psi ( three passages) . The extract was centrifuged at 16000 g for 30 minutes at 4°C.

**Analysis on Coomassie stained SDS-polyacrylamide gels and Western blots**

5 After centrifugation of extracts described above, aliquots of supernatant and pellet were analysed by SDS-polyacrylamide gel electrophoresis and Western blotting. A major band of about 31 kDa, localized in the pellet fraction, was visualised by Coomassie stained gels and identified in Western blots by rabbit polyclonal 22 J 70 anti-protein D and by monoclonal Penta-His (Qiagen cat. n° 34660) which detects 10 accessible histidine tail. The level of expression represents about 3 to 5 % of total protein.

**EXAMPLE XI: Construction of an *E. coli* strain expressing fusion Protein-D1/3-E6-his / HPV18**

**1. Construction of expression plasmid**

15 a) Plasmid pMG MCS prot D1/3 (= pRIT14589) is a derivative of pMG81 (described supra) in which the codons 4-81 of NS1 coding region from Influenza were replaced by the codons corresponding to residues Ser 20 → Thr 127 of mature protein D of Haemophilus Influenzae strain 772, biotype 2 (H. Janson *et al.*, 1991, Infection and Immunity, Jan. p.119-125). The sequence of Prot-D1/3 is followed by a multiple 20 cloning site (11 residues) and a coding region for a C-terminal histidine tail (6 His). This plasmid is used to express the fusion protein D1/3-E6-his.

HPV genomic E6 and E7 sequences type HPV18 (Cole *et al.*, J. Mol. Biol. 1987, 193 , p.599-608. ) were amplified from HPV18 full length genome cloned in pBR322 (obtained from Deutsches Krebsforschungszentrum (DKFZ), Referenzzentrum für 25 human pathogen Papillomaviruses - D 69120 - Heidelberg) and were subcloned into pUC19 to give TCA 302 (= pRIT14467).

**Construction of plasmid TCA 314(= pRIT14526) : a plasmid expressing the fusion Protein-D1/3-E6-His /HPV18**

The nucleotides sequences corresponding to amino acids

30 1 → 158 of E6 protein were amplified from pRIT14467. During the polymerase chain reaction, NcoI and SpeI restriction sites were generated at the 5' and 3' ends of the E6 sequences allowing insertion into the same sites of plasmid

pMGMCS Prot D1/3 to give plasmid TCA314 (= pRIT14526). The insert was sequenced to verify that no modification had been generated during the polymerase chain reaction. The protein and coding sequence for the fusion protein-D1/3-E6-His is described in sequence ID No. 20 and 21.

5      **Transformation of AR58 strain**

Plasmid pRIT14526 was introduced into *E. coli* AR58 (Mott et al., 1985, Proc. Natl. Acad. Sci., 82:88) a defective  $\lambda$  lysogen containing a thermosensitive repressor of the  $\lambda$  pL promoter.

3. **Growth and induction of bacterial strain - Expression of Prot-D1/3-E6-His**

10     Cells of AR58 transformed with plasmid pRIT14526 were grown in 100 ml of LB medium supplemented with 50  $\mu$ gr/ml of Kanamycin at 30°C. During the logarithmic phase of growth bacteria were shifted to 39°C to inactivate the  $\lambda$  repressor and turn on the synthesis of protein D1/3-E6-his. The incubation at 39°C was continued for 4 hours. Bacteria were pelleted and stored at -20C.

15     **4. Characterization of fusion Protein D1/3-E6-his**

Frozen cells are thawed and resuspended in 10 ml of PBS buffer. Cells are broken in a French pressure cell press SLM Aminco at 20.000 psi (three passages). The extract is centrifuged at 16.000 g for 30 minutes at 4°C. After centrifugation of extracts described above, aliquots of supernatant and pellet were analysed by SDS-20 polyacrylamide gel electrophoresis and Western blotting.

A major band of about 32 kDa, localized in the pellet fraction, was visualised by Coomassie stained gels and identified in Western blots by rabbit polyclonal anti-protein-D and by Ni-NTA conjugate coupled to calf intestinal alkaline phosphatase (Qiagen cat. n° 34510) which detects accessible histidine tail. The level of expression represents about 3-5 % of total protein.

25     **EXAMPLE XII: Construction of an *E. coli* strain expressing fusion Protein-D1/3-E6E7-his / HPV18**

*I. Construction of expression plasmid*

a) Plasmid pMG MCS prot D1/3 (= pRIT14589) is a derivative of pMG81 (described 30 supra) in which the codons 4-81 of NS1 coding region from Influenza were replaced by the codons corresponding to residues Ser 20 → Thr 127 of mature protein D of *Haemophilus Influenzae* strain 772, biotype 2 (H. Janson et al., 1991, Infection and

Immunity, Jan. p.119-125). The sequence of Prot-D1/3 is followed by a multiple cloning site (11 residues) and a coding region for a C-terminal histidine tail (6 His). This plasmid is used to express the fusion protein D1/3-E6E7-his.

- b) HPV genomic E6 and E7 sequences type HPV18 (Cole et al., J.Mol.Biol. 1987, 193, 599-608) were amplified from HPV18 full length genome cloned in pBR322 (obtained from Deutsches Krebsforschungszentrum (DKFZ), Referenzzentrum für human pathogen Papillomaviruses - D 69120 - Heidelberg) and were subcloned into pUC19 to give TCA 302 (= pRIT14467).
- c) The coding sequences for E6 and E7 in TCA302 (= pRIT 14467) were modified with a synthetic oligonucleotides adaptor (inserted between Hga I and Nsi I sites) introducing a deletion of 11 nucleotides between E6 and E7 genes, removing the stop codon of E6 and creating fused E6 and E7 coding sequences in the plasmid TCA320( = pRIT 14618 ).

**Construction of plasmid TCA 328(= pRIT14567) : a plasmid expressing the fusion Protein-D1/3-E6E7-His /HPV18**

The nucleotides sequences corresponding to amino acids

1 → 263 of fused E6E7 protein were amplified from pRIT14618. During the polymerase chain reaction, NcoI and SpeI restriction sites were generated at the 5' and 3' ends of the E6E7 fused sequences allowing insertion into the same sites of plasmid pMGMCS Prot D1/3 to give plasmid TCA328 (= pRIT14567). The insert was sequenced to verify that no modification had been generated during the polymerase chain reaction. The protein and coding sequence for the fusion protein-D1/3-E6E7-His is described in sequence ID No. 22 and 23.

**2. Transformation of AR58 strain**

25 Plasmid pRIT14567 was introduced into *E. coli* AR58 (Mott et al., 1985, Proc. Natl. Acad. Sci., 82:88) a defective λ lysogen containing a thermosensitive repressor of the λ pL promoter.

**3. Growth and induction of bacterial strain - Expression of Prot-D1/3-E6E7-His**

Cells of AR58 transformed with plasmid pRIT14512 were grown in 100 ml of 30 LB medium supplemented with 50 µgr/ml of Kanamycin at 30°C. During the logarithmic phase of growth bacteria were shifted to 39°C to inactivate the λ repressor

and turn on the synthesis of protein D1/3-E6E7-his. The incubation at 39°C was continued for 4 hours. Bacteria were pelleted and stored at -20C.

#### 4.Characterization of fusion Protein D1/3-E6E7-his

Frozen cells are thawed and resuspended in 10 ml of PBS buffer. Cells are  
5 broken in a French pressure cell press SLM Aminco at 20.000 psi (three passages).  
The extract is centrifuged at 16.000 g for 30 minutes at 4°C.

After centrifugation of extracts described above, aliquots of supernatant and pellet were analysed by SDS-polyacrylamide gel electrophoresis and Western blotting.

10 A major band of about 48 kDa, localized in the pellet fraction, was visualised by Coomassie stained gels and identified in Western blots by rabbit polyclonal anti-protein-D and by Ni-NTA conjugate coupled to calf intestinal alkaline phosphatase (Qiagen cat. n° 34510) which detects accessible histidine tail. The level of expression represents about 1 % of total protein.

#### 15 EXAMPLE XIII

The therapeutic potential of vaccine containing the PD1/3 E7 fusion protein and different CpG oligonucleotides were evaluated in the TC1 (E7 expressing tumour model.)

##### 1. Therapeutic experiments: protocol

20 10e6 TC1 cells, E7 expressing tumour cells: were injected subcutaneously (200µl) in the flank of C57BL/6 immunocompetent mice. Mice were vaccinated 7 and 14 days after the tumour challenge, with 5µg ProtD 1/3 E7 HPV16 injected intra-footpad (100µl : 50µl / footpad) in the presence of different adjuvants:

25 2 and 4 weeks after the second immunisation, 5 mice/group were killed and spleens or popliteal lymph nodes were taken and analyzed for immune response.

##### 1.2 Results

###### Groups of mice

- 1) PBS
- 2) ProtD1/3 E7 HPV16
- 30 3) ProtD1/3 E7 HPV16 + oligo 1: 1826 (WD 1001): TCC ATG ACG TTC CTG ACG TT
- 4) Oligo 1
- 5) ProtD1/3 E7 HPV16 + oligo 2/ 1758 (WD1002): TCT CCC AGC GTG CGC CAT

6) Oligo 2

**Tumour Growth;**

was monitored by measuring individual tumours twice a week.

Figure 1 : represents the mean tumour growth (in mm<sup>2</sup>)/group n=10 followed  
5 over 4 weeks.

- The injection of 10e6 TC1 cells injected subcutaneously give rise to a growing tumour in 100% of the animals.
- Vaccinating with ProtD1/3E7 or adjuvant alone: 100% of the animals develop a tumour.
- 10 • As shown in figure 1 and 2, in the groups of mice that received the antigen with a CpG oligonucleotide the mean tumour growth remained very low and very similar between groups, reflecting that the tumour growth either was slowed down or that several tumours were completely rejected.

The analysis of individual tumour growth 2 and 4 weeks after the latest  
15 vaccination showed that complete rejection in the groups were:

	Day 28 (n=10)	day 42 (n=5)
E7+oligo1 (1826)	40%	40%
Oligo1	0%	0%
E7+oligo2 (1758)	70%	40%
Oligo2	0%	0%

The mean tumour growth/group of mice vaccinated with PD1/3 E7+ the CpG oligos are quite similar and analysis of the individual tumour growth showed that the  
20 CpG oligos induce prolonged complete tumour rejection.

**Conclusion**

Both CpG (Oligo 2>oligo 1) induced complete tumour regression .

Lymphoproliferative response was analysed by *in vitro* restimulation of spleen and lymph nodes cells for 72 hrs with either PD1/3E7, the protein E7(Bollen)  
25 and PD (whole) PD1/3 (coated or not on latex μbeads) (10, 1, 0.1 μg/ml) 2 and 4 weeks post II.

- Positive controls (ConA stimulation) were positive.

- Surprisingly, no E7 specific and no PD specific proliferative response could be observed starting with spleen cells 2 or 4 weeks post II (probably due to a technical problem: data not shown).
- On the contrary, lymph node cells from mice that received ProtD1/3 E7 in CpG oligos 1 and 2 showed a very good E7 specific proliferative response although almost no PD (whole) specific response could be observed even at the highest concentration of 100µg/ml no PD1/3 specific responses was observed even when coated on latex µbeads.

Similar data were obtained 4 weeks post II.

#### 10 Serology

The anti E7 antibody response: IgG tot and isotypes (IgG1, IgG2a, IgG2b, IgGTot) were measured by ELISA using the E7 protein as coating antigen as described in the Materials and Methods. Figures 3 and 4 show the relative percentage of the different IgG isotypes in the total of IgGs, 2 and 4 weeks post II respectively.

- 15 The Oligos affect only weakly (oligo 2) or not at all (Oligo 1) the weak antibody response observed when PD1/3E7 alone was injected.
- The predominant E7 specific antibody subclass was clearly IgG2b for all the formulation tested (80-90% of the total IgGs).

The same results were obtained 4 weeks post II

#### 20 Isotypic profile of anti E7 responses (post II, pooled sera) exp. 97293

Groups	IgG1	IgG2a	IgG2b	IgGtot
1) PBS	0	0	0	0
2) ProtD1/3 E7 HPV16	1020	0	4130	4740
3) ProtD1/3 E7 HPV16 + oligo 1	170	400	3680	4910
4) Oligo 1	0	0	530	420
5) ProtD1/3 E7 HPV16 + oligo 2	0	590	7560	13690
6) Oligo 2	0	0	0	0

Groups	IgG1	IgG2a	IgG2b	IgGtot
1) PBS	0	0	0	0
2) ProtD1/3 E7 HPV1	240	0	1650	1400
3) ProtD1/3 E7 HPV16 + oligo 1	0	0	1280	1430
4) Oligo 1	0	0	0	0
5) ProtD1/3 E7 HPV16 + oligo 2	0	560	3600	5880
6) Oligo 2	0	0	0	0

**CTL assay:**

- A CTL response could be detected when measured 2 weeks after the latest vaccination, when cells were re-stimulated in vitro with irradiated TC1 when TC1 or peptide E7 pulsed EL4, were used as target cells, when mice immunised with PD1/3 E7 + CpG oligo 2> 1 (25-40% specific lysis) and not with oligos alone.
- 5 • Lysis was seen on TC1 cells than on peptide E7 pulsed EL4 cells, but this is mostly observed in the groups of mice vaccinated with PD1/3E7 + CpG oligos (2>1). In this experiment other formulations did not induce a CTL.
- 10 • Using E7 pulsed EL4 cells, no lysis was observed when mice received the protein or the adjuvant alone.

**1.3 Materials and Methods**

Component	Brand	Batch number	Concentration (mg/ml)	Buffer
ProtD1/3-E7		957/015	0.677	PBS 7.4
oligo CpG 1826	EuroGentec	WD1001	5	H <sub>2</sub> O
oligo CpG	EuroGentec	WD1002	5	H <sub>2</sub> O

**1.3.1 Formulation Process**

15 All the formulations were prepared on the day of injection.

**Oligo containing formulations**

Formulations containing oligo alone without other adjuvant were prepared by addition of CpG to the diluted PrtD1/3-E7 in PBS pH 7.4.

20 The adjuvant controls without antigen were prepared by replacing the protein by PBS.

**1.3.2 Mice and Cell lines**

Mice C57Bl/6 (Iffa Credo) 6-8 weeks old mice were used in these experiments.

Cell lines: TC1 (obtained from the John Hopkin's University), or EL4 cells were grown in RPMI 1640 (Bio Whittaker) containig 10% FCS and additives: 2mM L-Glutamine , 1% antibiotics (10000U/ml penicilin, 10000µg/ml streptomycin) 1% non essential amino acid 100x, 1% sodium pyruvate (Gibco), 5 10e-5 M 2-

mercaptoethanol. Before injection TC1 cells were trypsynized and washed in serum free medium.

### 1.3.3 Tumour growth:

All the animals were injected with tumor cells on day 0 and were randomized  
5 at day 7. Individual tumor growth was followed over time ( the 2 main diameters ( A,  
B) were measured using calipers twice a week, A x B represents the "tumor surface"  
and the average of the 5 values / groups is showed on a graphic over time: 6 weeks

### 1.3.4 CMI read out

#### In vitro lymphoproliferation

10 Lymphoproliferation was performed on individual spleens and on lymph node pools. 200000 spleen cells or popliteal lymph node cells were plated in triplicate, in 96 well microplate, in RPMI medium containing 1% normal mice serum and additives . After 72 hrs of in vitro re-stimulation with different amounts of PD1/3 E7 (1, 0.1, 0.01 µg/ml) or E7 ( 10-1-0.1 µg/ml ) After 72hrs, 100 µl of culture  
15 supernatant were removed and replaced by fresh medium containing 1µCi 3H thymidine (Amersham 5Ci/mmol) . After 16 hrs, cells were harvested onto filter plates. Incorporated radioactivity was counted in a β counter. Results are expressed in CPM (mean of triplicate wells) or as stimulation indexes (mean CPM in cultures with antigen / mean CPM in cultures without antigen).

20

### 1.3.5 CTL assay

20 10e6 spleen cells were co-cultured with 2 10e6 irradiated (18000r) TC1 cells (E7 expressing tumor) for 7 days in the presenced or absence of ConA sup. (2%)

25 Target cells used to assess cytotoxicity were either Cr51 ( DuPont NEN 37MBq/ml) loaded (1hr at 37°C) TC1 cells or E7 pulsed EL4 cells (for 1 hr at 37°C during the Cr 51 loading of the cells 10µg/ml of E7-derived peptide (49-57) (QCB) compared to EL4 cells NK dependant lysis was assessed on K562 target cells 2000 target cells were added / well of 96 well plate ( V bottom nunc 2-45128) with 100/1  
30 being the highest Effector / target ratio. Controls for spontaneous or maximal Cr51 release were performed in sextuplet and were targets in medium or in triton 1.5%. All plates were gently centrifuged and incubated for 4 hrs at 37 in 7% CO2. 50 µl of the

supernatant was deposited on 96w Lumaplate (Packard) let dry O/N and counted in a Top Count counter. Data is expressed as percent specific lysis which is calculated from the c.p.m. by the formula (experimental release - spontaneous release) / (maximal release - spontaneous release) X 100.

5                   **Serology**

Quantitation of anti E7 antibody was performed by Elisa using E7as coating antigen. Antigen and antibody solutions were used at 50 µl per well. Antigen was diluted at a final concentration of 3 µg/ml in carbonate buffer ph9.5 and was adsorbed overnight at 4°C to the wells of 96 wells microtiter plates (Maxisorb Immuno-plate, Nunc, Denmark). The plates were then incubated for 1hr at 37°C with PBS containing 1% bovine serum albumin and 0.1% Tween 20 (saturation buffer). Two-fold dilutions of sera (starting at 1/100 dilution) in the saturation buffer were added to the E7-coated plates and incubated for 1 hr 30 min at 37°C. The plates were washed 3 times with PBS 0.1% Tween 20 and biotin-conjugated anti-mouse IgG1, IgG2a or IgG2b or IgGtot (Amersham, UK) diluted 1/5000 in saturation buffer was added to each well and incubated for 1 hr 30 min at 37°C. After a washing step, streptavidin-biotinylated peroxidase complex (Amersham, UK) diluted 1/5000 in saturation buffer was added for an additional 30 min at 37°C. Plates were washed as above and incubated for 10 min with TMB( tetra-methyl-benzidine). The reaction was stopped with H<sub>2</sub>SO<sub>4</sub> 4N and read at 450 nm. Midpoint dilutions were calculated by SoftmaxPro (using a four parameters equation ).

10                  **EXAMPLE XIV**

In a second experiment, the vaccine of the invention were tested to assess the significance of the backbone:

15                  **Therapeutic experiment: protocol**

- 10e6 TC1 cells , E7 expressing tumor cells : were injected subcutaneously (200µl) in the flank of immunocompetent C57BL/6 mice.
- 30                  • 2 vaccinations, 7 and 14 days after the tumor challenge, with 5µg ProtD 1/3 E7 HPV16 injected intra- footpad (100 µl : 50µl / footpad) +/- CpG oligo; Oligo 1 (WD1001) as a phosphorothioate modified or the same Oligo (WD1006) but with phosphodiester linkage.

- 5 animals /group.

The tumor growth was monitored by measuring individual tumors twice a week and  
5 the mean tumor growth/ group of 5 animals is depicted in figure 5 and show the  
phosphorothioate modified oligonucleotides are effective in bringing about tumour  
regression.

#### **Conclusions:**

10

- All the animals that received 10e6 TC1 tumor cells develop a growing tumor.
- 100% of the animals vaccinated twice, 7 days apart, with the PD1/3 E7 HPV16 protein alone develop a tumor.
- 15 • 100% of the animals receiving the PD1/3 E7 protein + oligo WD1006 develop a tumor at the concentrations tested
- All the groups of animals that received the E7 protein + CpG 1001 at a concentration ranging from 10 to 200 $\mu$ g show tumor regression partial or  
20 complete(20-40%).

The first concentration at which this therapeutic effect on tumor regression is not fully obtained is E7+ 1 $\mu$ g CpG oligo 1001.

25 **EXAMPLE XV**

In a third series of experiments, the vaccines of the invention were evaluated in transgenic mice expressing E7 protein.

- 30
- The transgenic mouse strain has been generated by M. Parmentier and C. Ledent at the IRIBHN (ULB). (Ref: PNAS (USA) 1990, 87; 6176-6180).
  - As transgenic mice live with the E7 HPV16 gene from birth, they are considered "tolerant" to this gene: E7 from HPV 16, in this situation is considered as a "self  
35 antigen".

- The expression of the transgene is driven by the thyroglobulin promoter. As Thyroglobulin is constitutively expressed only in the Thyroid, E7 is expressed in the thyroid.
- 5     • As a consequence of this expression, thyroid cells proliferate, mouse develop goiter and nodules which after 6 months to 1 year can evolve in invasive cancer.

The results (figure 6) of the experiments show that therapeutic vaccination with CpG oligonucleotide and antigen as described herein, results in a reduction of tumour growth and can induce complete tumour regression.  
10

### **Material & Methods**

- 10e6 TC1 cells, E7 expressing tumor cells : were injected subcutaneously (200µl) in the flank of male or female C57BL/6 Transgenic
- 15     • mice were vaccinated 7 and 14 days after the tumor challenge, with 5µg ProtD 1/3 E7 HPV16 injected intra- footpad (100 µl : 50µl / footpad) in the 2 presence of CpG oligonucleotide TCT CCC AGC GTG CGC CAT and two control  
20     adjuvants:,
  - 10 animals /group
- 25     2 and 4 weeks after the second immunization were killed and spleens or popliteal lymph.

### **Conclusion**

The vaccines of the invention are effective in bringing about tumour regression in  
30     HPV induced tumours.

**CLAIMS**

1. A composition comprising an E6 or E7 protein or E6/E7 fusion protein from HPV optionally linked to an immunological fusion partner, and an immunomodulatory CpG oligonucleotide.  
5
2. A composition as claimed in claim 1 wherein the fusion partner is selected from the group; protein D or a fragment thereof from *Haemophilus influenzae* B, lipoprotein D or fragment thereof from *Haemophilus influenzae* B, NS1 or fragment thereof from *Influenzae Virus*, and LYTA or fragment thereof from  
10 *Streptococcus Pneumoniae*.
3. A composition as claimed in claim 1 or 2 wherein the E6 or E7 proteins are derived from HPV16 or HPV18.
4. A composition as claimed in claim 1, 2 or 3 wherein the E7 protein is mutated.
5. A composition as claimed in claim 1, 2 or 3 wherein the E6 protein is mutated.  
15
6. A composition as claimed in any of claims 1 to 5 additionally comprising a histidine tag of at least 4 histidine residues.
7. A composition as claimed herein comprising an additional HPV antigen.
8. A composition as claimed herein where the immunomodulatory CpG oligonucleotide comprises a hexamer motif: purine purine cytosine guanine pyrimidine pyrimidine.  
20
9. A composition as claimed herein wherein the immunomodulatory CpG oligonucleotide has two or more CpG motifs.
10. A composition as claimed herein wherein the CpG oligonucleotide contains a phosphorothioate inter-nucleotide linkage.  
25
11. A composition as claimed herein wherein the CpG oligonucleotide is selected from the group:

OLIGO 1: TCC ATG ACG TTC CTG ACG TT

OLIGO 2: TCT CCC AGC GTG CGC CAT

OLIGO 3: ACC GAT GAC GTC GCC GGT GAC GGC ACC ACG

12. A composition as claimed herein for use in medicine.
- 5      13. A method of inducing an immune response in a patient to an HPV antigen comprising administering a safe and effective amount of a composition as claimed herein.
14. A method of preventing or treating HPV induced tumours in a patient comprising administering a safe and effective amount of a composition as claimed herein.
- 10     15. A method of preparing a composition as claimed herein, comprising admixing an E6, E7 or E6/E7 fusion protein optionally linked to an immunological fusion partner, and an immunomodulatory CpG oligonucleotide.

Fig. 1

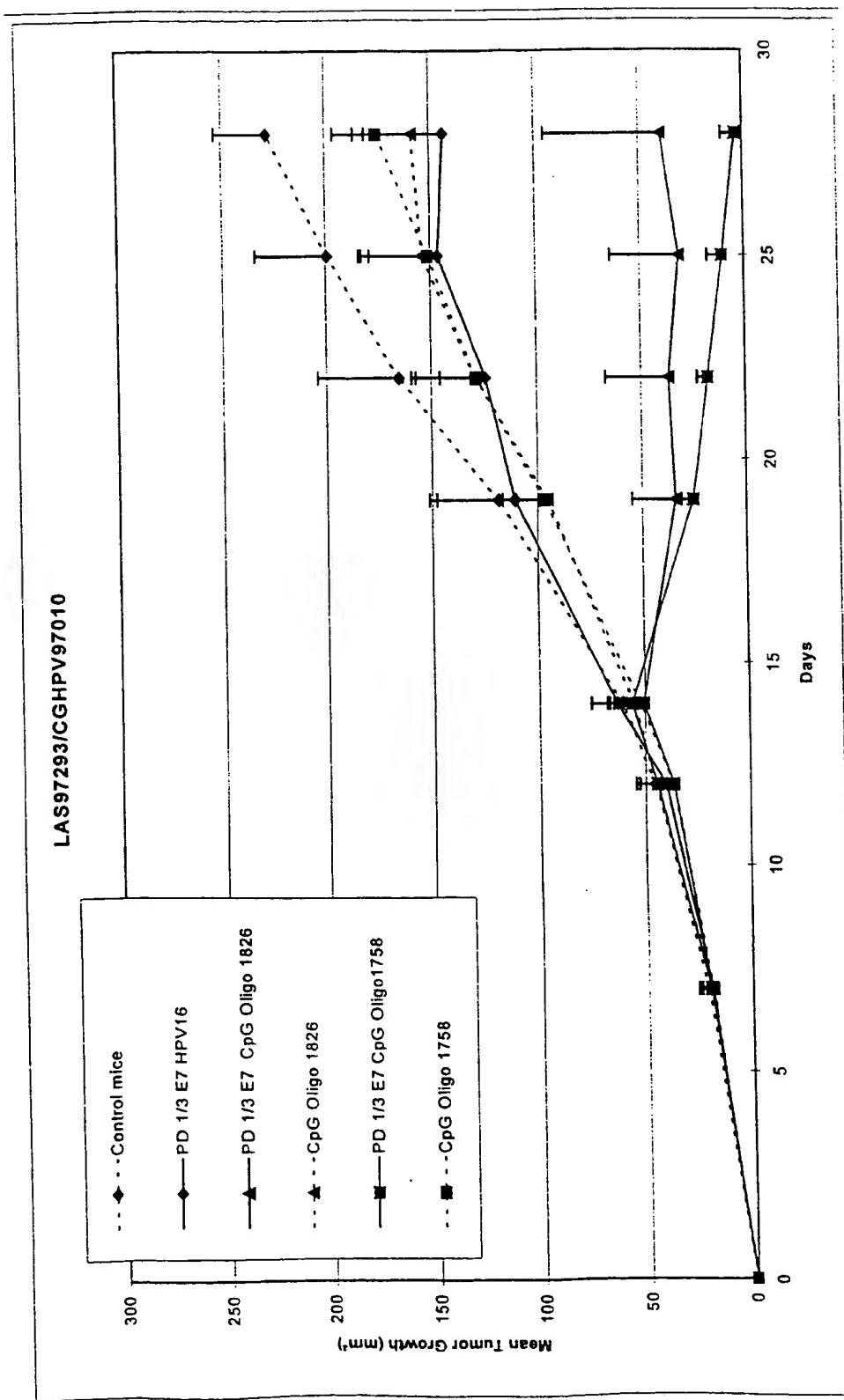


Fig. 2

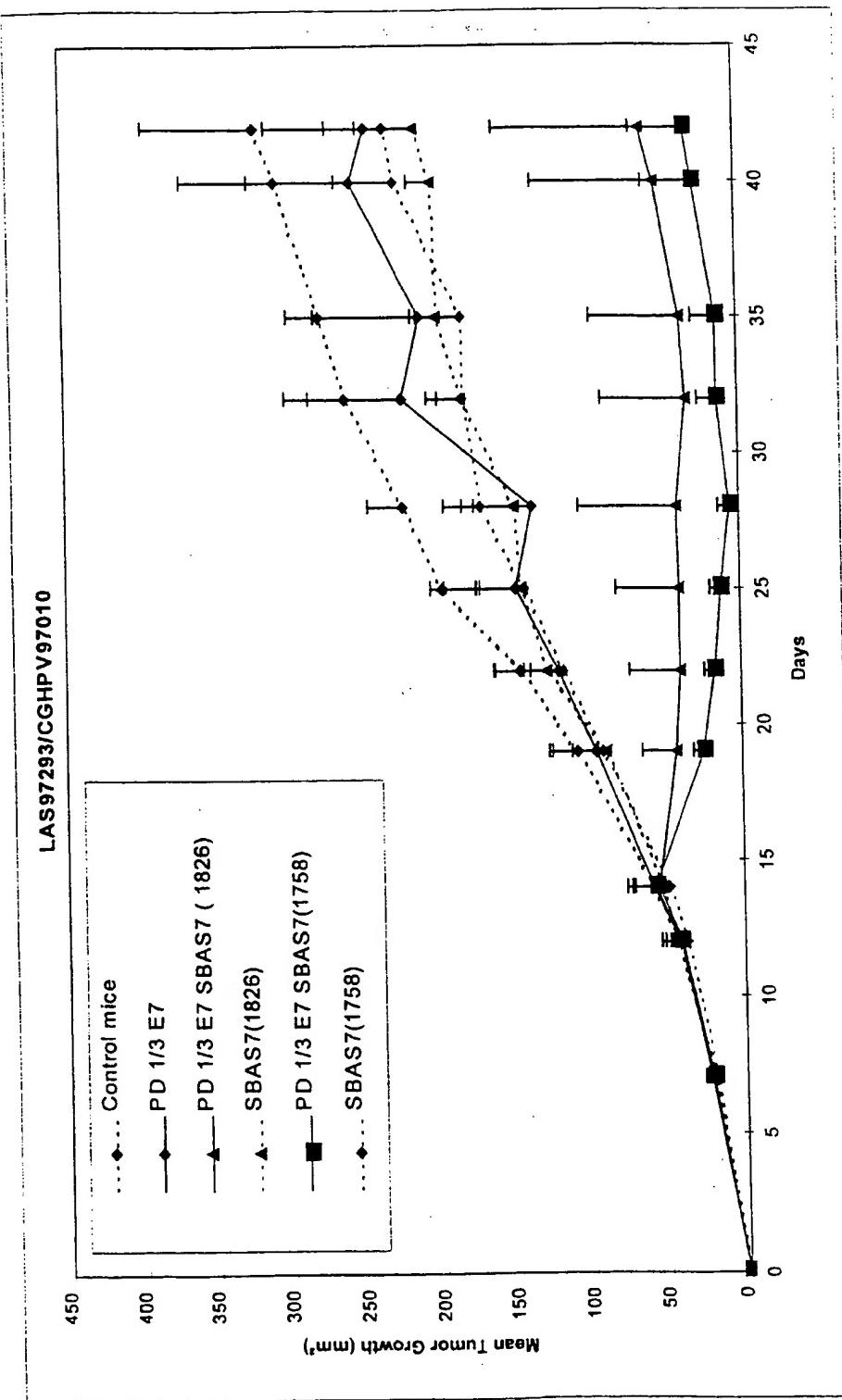


Fig. 3

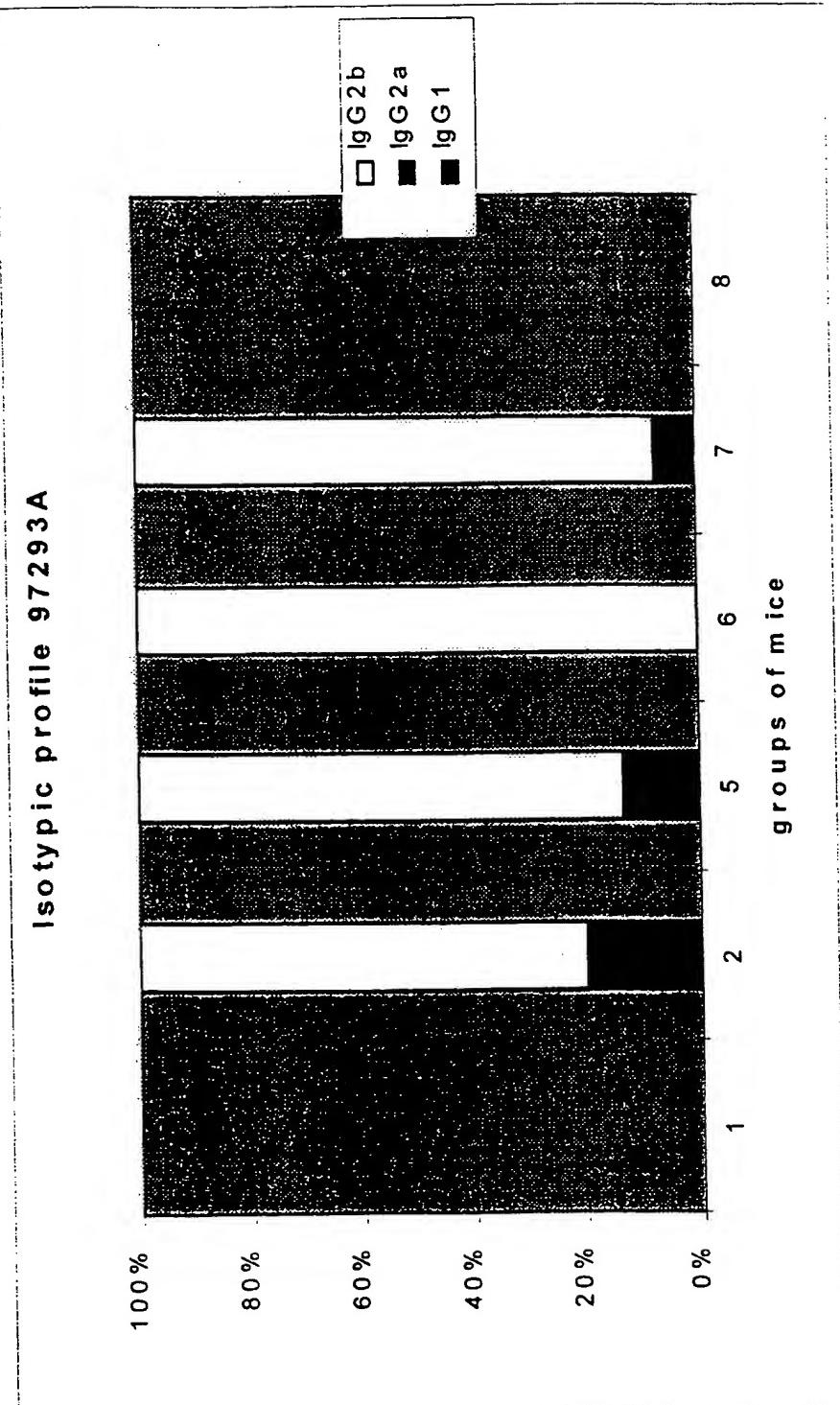


Fig. 4

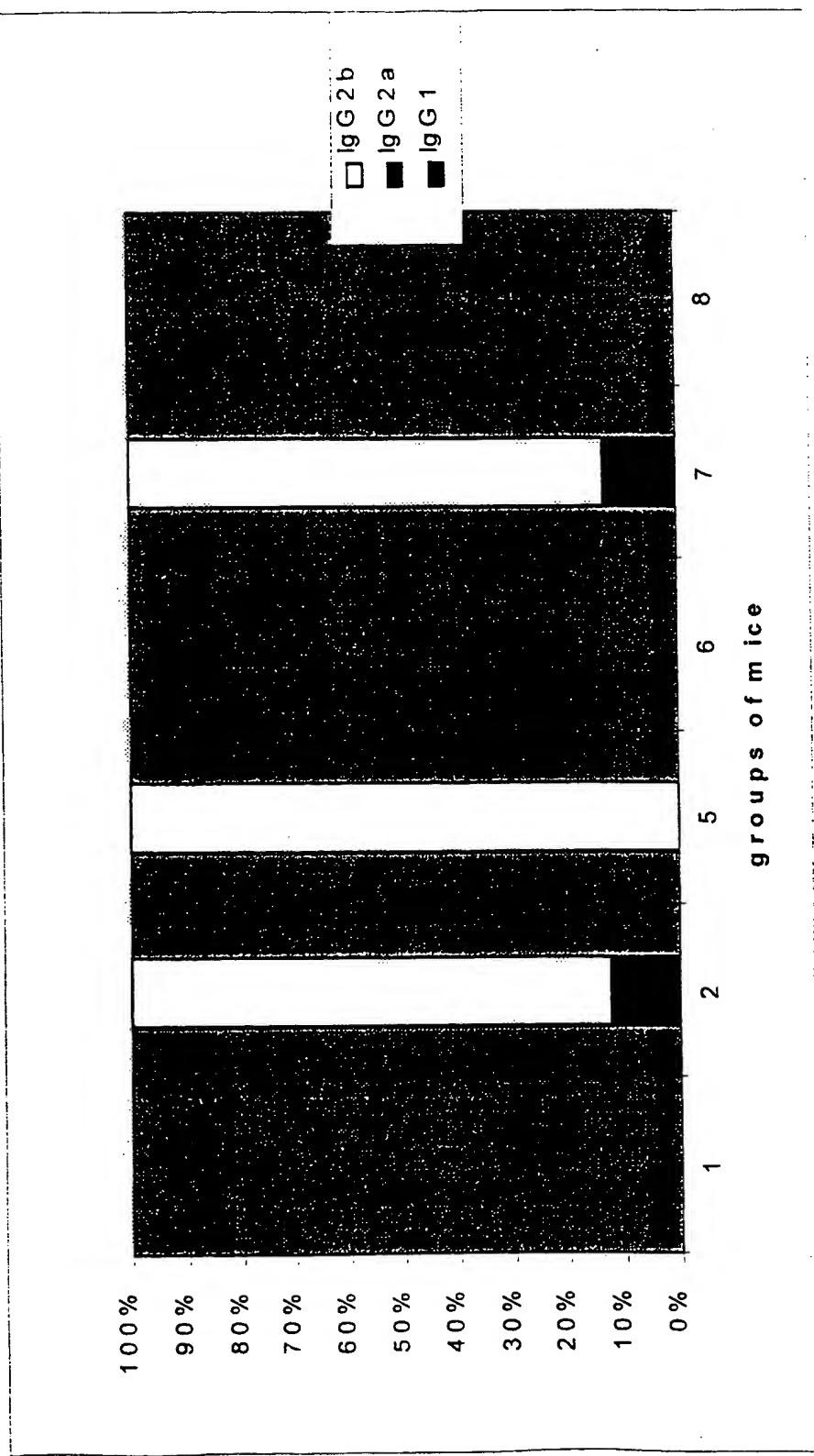


Fig. 5

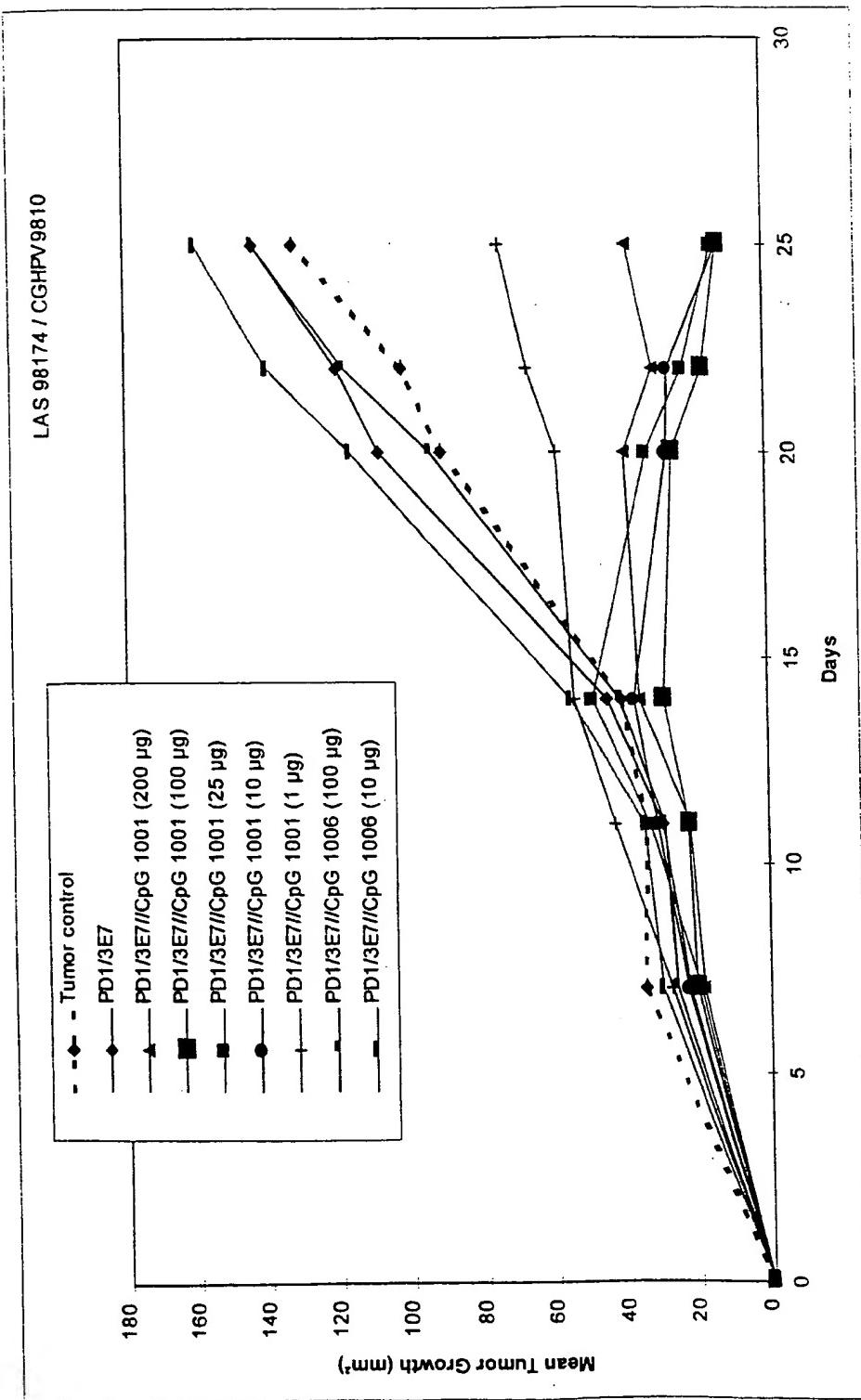
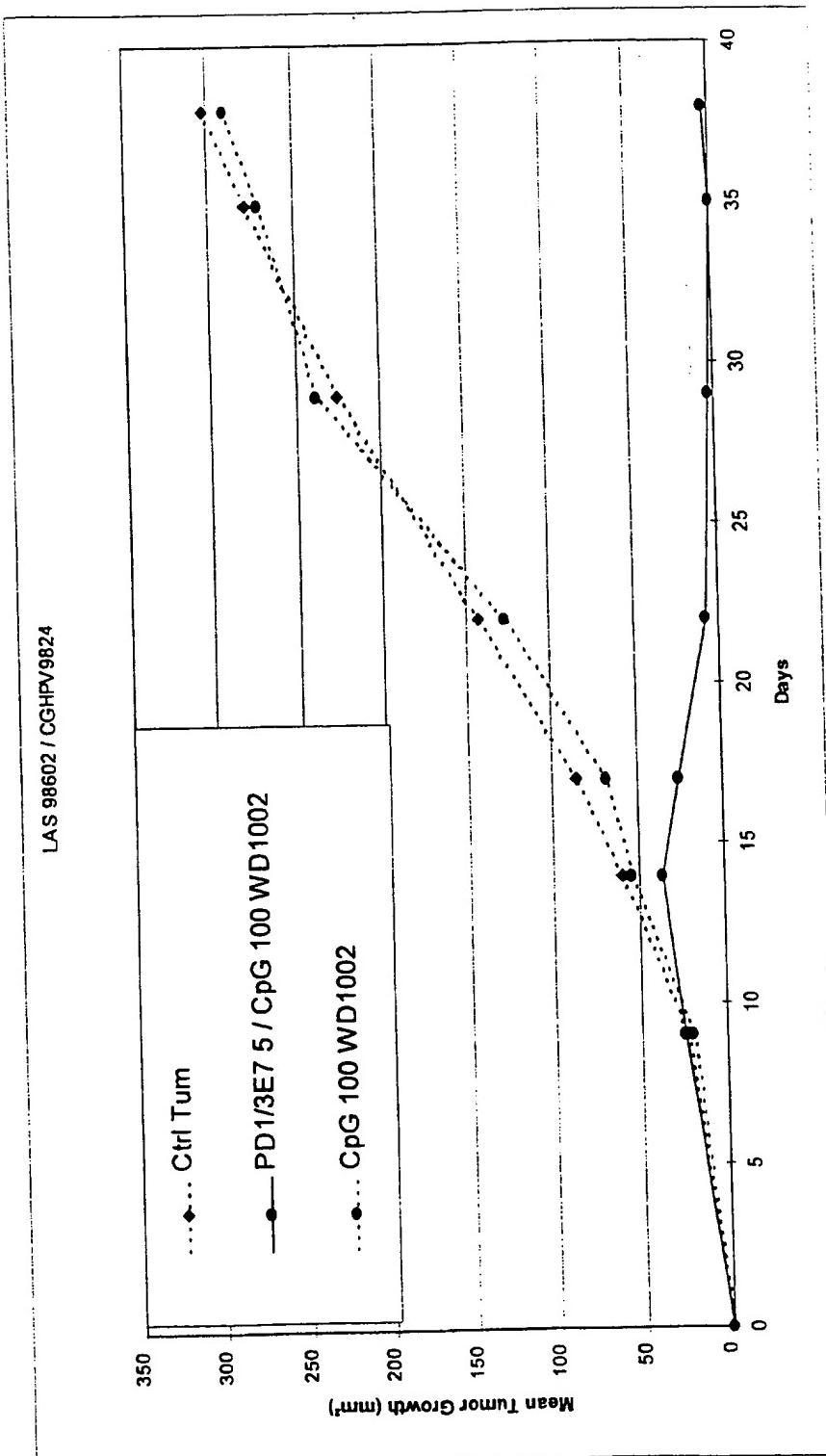


Fig. 6



## SEQUENCE LISTING

## (1) GENERAL INFORMATION

5

(i) APPLICANT: BRUCK, CLAUDINE

(ii) TITLE OF THE INVENTION: VACCINE

10

(iii) NUMBER OF SEQUENCES: 23

15

(iv) CORRESPONDENCE ADDRESS:

- (A) ADDRESSEE: SmithKline Beecham
- (B) STREET: 2 New Horizons Court, Great West Road, B
- (C) CITY: Middx
- (D) STATE:
- (E) COUNTRY: UK
- (F) ZIP: TW8 9EP

20

(v) COMPUTER READABLE FORM:

- (A) MEDIUM TYPE: Diskette
- (B) COMPUTER: IBM Compatible
- (C) OPERATING SYSTEM: DOS

25

(D) SOFTWARE: FastSEQ for Windows Version 2.0

30

(vi) CURRENT APPLICATION DATA:

- (A) APPLICATION NUMBER:
- (B) FILING DATE:

35

(vii) PRIOR APPLICATION DATA:

- (A) APPLICATION NUMBER:
- (B) FILING DATE:

40

(viii) ATTORNEY/AGENT INFORMATION:

- (A) NAME: Dalton, Marcus J
- (B) REGISTRATION NUMBER:
- (C) REFERENCE/DOCKET NUMBER: B45124

45

(ix) TELECOMMUNICATION INFORMATION:

- (A) TELEPHONE: 0181 9756348
- (B) TELEFAX: 0181 9756177
- (C) TELEX:

50

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 220 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear  
Protein D 1/3 E7 his

55

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

60

Met	Asp	Pro	Ser	Ser	His	Ser	Ser	Asn	Met	Ala	Asn	Thr	Gln	Met	Lys
1					5				10					15	
Ser	Asp	Lys	Ile	Ile	Ile	Ala	His	Arg	Gly	Ala	Ser	Gly	Tyr	Leu	Pro
					20			25					30		
Glu	His	Thr	Leu	Glu	Ser	Lys	Ala	Leu	Ala	Phe	Ala	Gln	Gln	Ala	Asp
					35			40				45			
Tyr															

65

Tyr	Leu	Glu	Gln	Asp	Leu	Ala	Met	Thr	Lys	Asp	Gly	Arg	Leu	Val	Val
-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

50	55	60
Ile His Asp His Phe Leu Asp Gly Leu Thr Asp Val Ala Lys Lys Phe		
65	70	75
Pro His Arg His Arg Lys Asp Gly Arg Tyr Tyr Val Ile Asp Phe Thr		80
5	85	90
Leu Lys Glu Ile Gln Ser Leu Glu Met Thr Glu Asn Phe Glu Thr Met		95
100	105	110
Ala Met His Gly Asp Thr Pro Thr Leu His Glu Tyr Met Leu Asp Leu		
115	120	125
10 Gln Pro Glu Thr Thr Asp Leu Tyr Cys Tyr Glu Gln Leu Asn Asp Ser		
130	135	140
Ser Glu Glu Glu Asp Glu Ile Asp Gly Pro Ala Gly Gln Ala Glu Pro		
145	150	155
Asp Arg Ala His Tyr Asn Ile Val Thr Phe Cys Cys Lys Cys Asp Ser		160
15 165		170
Thr Leu Arg Leu Cys Val Gln Ser Thr His Val Asp Ile Arg Thr Leu		175
180	185	190
Glu Asp Leu Leu Met Gly Thr Leu Gly Ile Val Cys Pro Ile Cys Ser		
195	200	205
20 Gln Lys Pro Thr Ser Gly His His His His His His		
210	215	220

## (2) INFORMATION FOR SEQ ID NO:2:

25 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 663 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear  
 30 Protein D 1/3 E7 his

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

35	ATGGATCCAA GCAGCCATTG ATCAAATATG GCGAATAACCC AAATGAAATC AGACAAAATC
60	ATTATTGCTC ACCGTGGTGC TAGCGTTAT TTACCAAGAGC ATACGTTAGA ATCTAAAGCA
120	CTTGCCTTTG CACAACAGGC TGATTATTTA GAGCAAGATT TAGCAATGAC TAAGGATGGT
180	CGTTTAGTGG TTATTCACGA TCACTTTTA GATGGCTTGA CTGATGTTGC GAAAAAATT
40 240	CCACATCGTC ATCGTAAAGA TGGCCGTTAC TATGTCATCG ACTTTACCTT AAAAGAAATT
300	CAAAGTTAG AAATGACAGA AAACTTGAA ACCATGGCCA TGCATGGAGA TACACCTACA
45 360	TTGCATGAAT ATATGTTAGA TTTGCAACCA GAGACAATG ATCTCTACTG TTATGAGCAA
420	TTAAATGACA GCTCAGAGGA GGAGGATGAA ATAGATGGTC CAGCTGGACA AGCAGAACCG
480	540 GACAGAGCCC ATTACAATAT TGTAACCTTT TGGTCAAGT GTGACTCTAC GCTTCGGTTG
50	TGGTACAAA GCACACACGT AGACATTCTG ACTTTGGAAG ACCTGTTAAT GGGCACACTA
600	660 GGAATTGTGT GCCCCATCTG TTCTCAGAAA CCAACTAGTG GCCACCACATCA CCATCACCAT
55	TAA
663	

## (2) INFORMATION FOR SEQ ID NO:3:

60 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 822 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear  
 65 Protein D 1/3 E6 His/HPV 16

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

5           ATGGATCCAA GCAGCCATTG ATCAAATATG GCGAATAACCC AAATGAAATC AGACAAAATC  
 60          ATTATTGCTC ACCGTGGTGC TAGCGTTAT TTACCAAGAGC ATACGTTAGA ATCTAAAGCA  
 120         CTTGCCTTG CACAACAGGC TGATTATTGAGA GAGCAAGATT TAGCAATGAC TAAGGATGGT  
 180         CGTTTAGTGG TTATTACAGA TCACTTTA GATGGCTTGA CTGATGTTGC GAAAAAAATC  
 240         CCACATCGTC ATCGTAAAGA TGGCCGTTAC TATGTCATCG ACTTTACCTT AAAAGAAATT  
 300         CAAAGTTAG AAATGACAGA AAACATTGAA ACCATGGCCA TGTTTCAGGA CCCACAGGAG  
 360         CGACCCAGAA AGTTACCACA GTTATGCACA GAGCTGCAAA CAACTATACA TGATATAATA  
 420         TTAGAACATGTG TGTACTGCAGA GCAACAGTTA CTGGCACGTG AGGTATATGA CTTTGCTTT  
 480         CGGGATTTAT GCATAGTATA TAGAGATGGG AATCCATATG CTGTATGTGA TAAATGTTA  
 540         AAGTTTTATT CTAAAATTAG TGAGTATAGA CATTATTGTT ATAGTTTGTA TGGAAACAACA  
 600         TTAGAACAGC AATACAACAA ACCGTTGTGT GATTGTTAA TTAGGTGTAT TAACTGTCAA  
 660         AAGCCACTGT GTCCTGAAGA AAAGCAAAGA CATCTGGACA AAAAGCAAAG ATTCCATAAT  
 720         ATAAGGGGTC GGTGGACCGG TCGATGTATG TCTTGTGCA GATCATCAAG AACACGTAGA  
 780         GAAACCCAGC TGACTAGTGG CCACCATCAC CATCACCATT AA  
 822

## (2) INFORMATION FOR SEQ ID NO:4:

35           (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 274 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear  
 40           Protein D 1/3 E6 His/HPV 16

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

45           Met Asp Pro Ser Ser His Ser Ser Asn Met Ala Asn Thr Gln Met Lys  
 1               5               10               15  
 Ser Asp Lys Ile Ile Ile Ala His Arg Gly Ala Ser Gly Tyr Leu Pro  
 20              25              30  
 Glu His Thr Leu Glu Ser Lys Ala Leu Ala Phe Ala Gln Gln Ala Asp  
 35              40              45  
 Tyr Leu Glu Gln Asp Leu Ala Met Thr Lys Asp Gly Arg Leu Val Val  
 50              55              60  
 Ile His Asp His Phe Leu Asp Gly Leu Thr Asp Val Ala Lys Lys Phe  
 65              70              75              80  
 Pro His Arg His Arg Lys Asp Gly Arg Tyr Tyr Val Ile Asp Phe Thr  
 85              90              95  
 Leu Lys Glu Ile Gln Ser Leu Glu Met Thr Glu Asn Phe Glu Thr Met  
 100            105            110  
 Ala Met Phe Gln Asp Pro Gln Glu Arg Pro Arg Lys Leu Pro Gln Leu  
 115            120            125  
 60            Cys Thr Glu Leu Gln Thr Thr Ile His Asp Ile Ile Leu Glu Cys Val  
 130            135            140  
 Tyr Cys Lys Gln Gln Leu Leu Arg Arg Glu Val Tyr Asp Phe Ala Phe  
 145            150            155              160  
 Arg Asp Leu Cys Ile Val Tyr Arg Asp Gly Asn Pro Tyr Ala Val Cys  
 165            170            175  
 65            Asp Lys Cys Leu Lys Phe Tyr Ser Lys Ile Ser Glu Tyr Arg His Tyr

180 Cys Tyr Ser Leu Tyr Gly Thr Thr 195 Leu Cys Asp Leu Leu Ile Arg Cys Ile Asn Cys Gln Lys Pro 210 Pro Glu Glu Lys Gln Arg His Leu Asp Lys Lys Gln Arg Phe His Asn 225 Ile Arg Gly Arg Trp Thr Gly Arg Cys Met Ser Cys Cys Arg Ser Ser 245 10 Arg Thr Arg Arg Glu Thr Gln Leu Thr Ser Gly His His His His His 260 His	185 Leu Glu Gln Gln Tyr Asn Lys Pro 200 205 215 220 230 235 240 250 255 265 270	190 195 200 205 210 215 220 225 230 235 240 245 250 255 260 265 270
---	---	---

## 15 (2) INFORMATION FOR SEQ ID NO:5:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1116 base pairs
- (B) TYPE: nucleic acid
- 20 (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

Protein D 1/3 E6/E7/ HPV16

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

25 ATGGATCCAA GCAGCCATTG ATCAAATATG GCGAATACCC AAATGAAATC AGACAAAATC  
 60 ATTATTGCTC ACCGTGGTGC TAGCGGTTAT TTACCAAGAGC ATACGTTAGA ATCTAAAGCA  
 120 CTTGCGTTTG CACAACAGGC TGATTATTTA GAGCAAGATT TAGCAATGAC TAAGGATGGT  
 180 CGTTTAGTGG TTATTACAGA TCACTTTTA GATGGCTTGA CTGATGTTGC GAAAAAAATTC  
 240 CCACATCGTC ATCGTAAAGA TGGCCGTTAC TATGTCATCG ACTTTACCTT AAAAGAAATT  
 300 300 CAAAGTTTAG AAATGACAGA AAACTTGAA ACCATGGCCA TGTTCAGGA CCCACAGGAG  
 360 CGACCCAGAA AGTTACCACA GTTATGCACA GAGCTGCAA CAACTATAACA TGATATAATA  
 420 420 TTAGAATGTG TGTACTGCAA GCAACAGTTA CTGCGACGTG AGGTATATGA CTTTGCTTT  
 480 CGGGATTTAT GCATAGTATA TAGAGATGGG AATCCATATG CTGTATGTGA TAAATGTTA  
 540 540 AAGTTTTATT CTAAAATTAG TGAGTATAGA CATTATTGTT ATAGTTGTA TGGAACAACA  
 600 600 TTAGAACAGC AATACAACAA ACCGTTGTGT GATTTGTTAA TTAGGTGTAT TAACTGTCAA  
 660 AAGCCACTGT GTCCTGAAGA AAAGCAAAGA CATCTGGACA AAAAGCAAAG ATTCCATAAT  
 720 720 ATAAGGGGTC GGTGGACCGG TCGATGTATG TCTTGTGCA GATCATCAAG AACACGTAGA  
 780 GAAACCCAGC TGATGCATGG AGATACACCT ACATTGCATG AATATATGTT AGATTTGCAA  
 840 CCAGAGACAA CTGATCTCA CTGTTATGAG CAATTAAATG ACAGCTCAGA GGAGGGAGGAT  
 900 900 GAAATAGATG GTCCAGCTGG ACAAGCAGAA CCGGACAGAG CCCATTACAA TATTGTAACC  
 960 960 TTTTGTGCA AGTGTGACTC TACGCTTCGG TTGTGCGTAC AAAGCACACA CGTAGACATT  
 1020 1020 CGTACTTTGG AAGACCTGTT AATGGGCACA CTAGGAATTG TGTGCCCAT CTGTTCTCAG  
 1080 1080 AAACCAACTA GTGGCCACCA TCACCATCAC CATTAA  
 1116

## 65 (2) INFORMATION FOR SEQ ID NO:6:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 372 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

5 Protein D 1/3 E6/E7/ HPV16

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

10	Met Asp Pro Ser Ser His Ser Ser Asn Met Ala Asn Thr Gln Met Lys	
	1 5 10 15	
	Ser Asp Lys Ile Ile Ala His Arg Gly Ala Ser Gly Tyr Leu Pro	
	20 25 30	
	Glu His Thr Leu Glu Ser Lys Ala Leu Ala Phe Ala Gln Gln Ala Asp	
	35 40 45	
15	Tyr Leu Glu Gln Asp Leu Ala Met Thr Lys Asp Gly Arg Leu Val Val	
	50 55 60	
	Ile His Asp His Phe Leu Asp Gly Leu Thr Asp Val Ala Lys Lys Phe	
	65 70 75 80	
20	Pro His Arg His Arg Lys Asp Gly Arg Tyr Tyr Val Ile Asp Phe Thr	
	85 90 95	
	Leu Lys Glu Ile Gln Ser Leu Glu Met Thr Glu Asn Phe Glu Thr Met	
	100 105 110	
	Ala Met Phe Gln Asp Pro Gln Glu Arg Pro Arg Lys Leu Pro Gln Leu	
	115 120 125	
25	Cys Thr Glu Leu Gln Thr Thr Ile His Asp Ile Ile Leu Glu Cys Val	
	130 135 140	
	Tyr Cys Lys Gln Gln Leu Leu Arg Arg Glu Val Tyr Asp Phe Ala Phe	
	145 150 155 160	
30	Arg Asp Leu Cys Ile Val Tyr Arg Asp Gly Asn Pro Tyr Ala Val Cys	
	165 170 175	
	Asp Lys Cys Leu Lys Phe Tyr Ser Lys Ile Ser Glu Tyr Arg His Tyr	
	180 185 190	
	Cys Tyr Ser Leu Tyr Gly Thr Thr Leu Glu Gln Gln Tyr Asn Lys Pro	
	195 200 205	
35	Leu Cys Asp Leu Leu Ile Arg Cys Ile Asn Cys Gln Lys Pro Leu Cys	
	210 215 220	
	Pro Glu Glu Lys Gln Arg His Leu Asp Lys Lys Gln Arg Phe His Asn	
	225 230 235 240	
40	Ile Arg Gly Arg Trp Thr Gly Arg Cys Met Ser Cys Cys Arg Ser Ser	
	245 250 255	
	Arg Thr Arg Arg Glu Thr Gln Leu Met His Gly Asp Thr Pro Thr Leu	
	260 265 270	
	His Glu Tyr Met Leu Asp Leu Gln Pro Glu Thr Thr Asp Leu Tyr Cys	
	275 280 285	
45	Tyr Glu Gln Leu Asn Asp Ser Ser Glu Glu Glu Asp Glu Ile Asp Gly	
	290 295 300	
	Pro Ala Gly Gln Ala Glu Pro Asp Arg Ala His Tyr Asn Ile Val Thr	
	305 310 315 320	
50	Phe Cys Cys Lys Cys Asp Ser Thr Leu Arg Leu Cys Val Gln Ser Thr	
	325 330 335	
	His Val Asp Ile Arg Thr Leu Glu Asp Leu Leu Met Gly Thr Leu Gly	
	340 345 350	
55	Ile Val Cys Pro Ile Cys Ser Gln Lys Pro Thr Ser Gly His His His	
	355 360 365	
	His His His	
	370	

## (2) INFORMATION FOR SEQ ID NO:7:

60

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 663 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

65

Protein D 1/3 E7 mutated HPV 16

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

ATGGATCCAA GCAGCCATTG ATCAAATATG GCGAATACCC AAATGAAATC AGACAAAATC  
5 60 ATTATTGCTC ACCGTGGTGC TAGCGTTAT TTACCAGAGC ATACGTAGA ATCTAAAGCA  
120 CTTGCGTTG CACAACAGGC TGATTATTTA GAGCAAGATT TAGCAATGAC TAAGGATGGT  
180 CGTTTAGTGG TTATTACGGA TCACTTTTA GATGGCTTGA CTGATGTTGC GAAAAAAATTC  
240 CCACATCGTC ATCGTAAAGA TGGCCGTTAC TATGTCATCG ACTTTACCTT AAAAGAAATT  
300 CAAAGTTTAG AAATGACAGA AAACTTGAA ACCATGGCCA TGCATGGAGA TACACCTACA  
360 TTGCATGAAT ATATGTTAGA TTTGCAACCA GAGACAAC TGACTCTACGG TTATCAGCAA  
420 TAAATGACA GCTCAGAGGA GGAGGATGAA ATAGATGGTC CAGCTGGACA ACCAGAACCG  
480 GACAGAGGCC ATTACAATAT TGTAACCTT TTGCAAGT GTGACTCTAC GCTTCGGTTG  
540 TGCACACACAA GCACACACGT AGACATTGCT ACCTTGGAAAG ACCTGTTAAT GGGCACACTA  
600 GGAATTGTGT GCCCCATCTG TTCTCAGAAA CCAACTAGTG GCCACCATCA CCATCACCCT  
25 660 TAA  
663

(2) INFORMATION FOR SEQ ID NO:8:

30 (i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 220 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear

35

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:  
Protein D 1/3 E7 mutated HPV 16

xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

40	Met Asp Pro Ser Ser His Ser Ser Asn Met Ala Asn Thr Gln Met Lys 1 5 10 15
	Ser Asp Lys Ile Ile Ile Ala His Arg Gly Ala Ser Gly Tyr Leu Pro 20 25 30
	Glu His Thr Leu Glu Ser Lys Ala Leu Ala Phe Ala Gln Gln Ala Asp 35 40 45
45	Tyr Leu Glu Gln Asp Leu Ala Met Thr Lys Asp Gly Arg Leu Val Val 50 55 60
	Ile His Asp His Phe Leu Asp Gly Leu Thr Asp Val Ala Lys Lys Phe 65 70 75 80
50	Pro His Arg His Arg Lys Asp Gly Arg Tyr Tyr Val Ile Asp Phe Thr 85 90 95
	Leu Lys Glu Ile Gln Ser Leu Glu Met Thr Glu Asn Phe Glu Thr Met 100 105 110
	Ala Met His Gly Asp Thr Pro Thr Leu His Glu Tyr Met Leu Asp Leu 115 120 125
55	Gln Pro Glu Thr Thr Asp Leu Tyr Gly Tyr Gin Gln Leu Asn Asp Ser 130 135 140
	Ser Glu Glu Glu Asp Glu Ile Asp Gly Pro Ala Gly Gln Ala Glu Pro 145 150 155 160
60	Asp Arg Ala His Tyr Asn Ile Val Thr Phe Cys Cys Lys Cys Asp Ser 165 170 175
	Thr Leu Arg Leu Cys Val Gln Ser Thr His Val Asp Ile Arg Thr Leu 180 185 190
	Glu Asp Leu Leu Met Gly Thr Leu Gly Ile Val Cys Pro Ile Cys Ser 195 200 205
65	Gln Lys Pro Thr Ser Gly His His His His His His His His His

210

215

220

## (2) INFORMATION FOR SEQ ID NO:9:

- 5                   (i) SEQUENCE CHARACTERISTICS:  
                   (A) LENGTH: 879 base pairs  
                   (B) TYPE: nucleic acid  
                   (C) STRANDEDNESS: single  
                   (D) TOPOLOGY: linear  
10                   CLYTA E6 His HPV 16

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

15                   ATGAAAGGGG GAATTGTACA TTCAGACGGC TCTTATCCAA AAGACAAGTT TGAGAAAATC  
        60                   AATGGCACTT GGTACTACTT TGACAGTTCA GGCTATATGC TTGCAGACCG CTGGAGGAAG  
        120                  CACACAGACG GCAACTGGTA CTGGTCGAC AACTCAGGCG AAATGGCTAC AGGCTGGAAG  
        180                  AAAATCGCTG ATAAGTGGTA CTATTCAAC GAAGAAGGTG CCATGAAGAC AGGCTGGTC  
        240                  AAGTACAAGG ACACTTGGTA CTACTTAGAC GCTAAAGAAG GCGCCATGGT ATCAAATGCC  
        300                  TTTATCCAGT CAGCGGACGG AACAGGCTGG TACTACCTCA AACCAGACGG AACACTGGCA  
        360                  GACAGGCCAG AATTGGCCAG CATGCTGGAC ATGGCCATGT TTCAGGACCC ACAGGAGCGA  
        420                  CCCAGAAAGT TACCACAGTT ATGCACAGAG CTGCAAACAA CTATACATGA TATAATATTA  
        480                  GAATGTGTGT ACTGCAAGCA ACAGTTACTG CGACGTGAGG TATATGACTT TGCTTTCGG  
        540                  GATTTATGCA TAGTATATAG AGATGGAAT CCATATGCTG TATGTGATAA ATGTTTAAAG  
        600                  TTTTATTCTA AAATTAGTGA GTATAGACAT TATTGTTATA GTTTGTATGG ACAACACATTA  
        660                  GAACAGCAAT ACAACAAACC GTTGTGTGAT TTGTTAATTA GGTGTATTAA CTGTCAAAAG  
        720                  CCACTGTGTC CTGAAGAAAA GCAAAGACAT CTGGACAAAA AGCAAAGATT CCATAATATA  
        780                  AGGGGTCGGT GGACCGGTG ATGTATGTCT TGGTGCAGAT CATCAAGAAC ACGTAGAGAA  
        840                  ACCCAGCTGA CTAGTGGCCA CCATCACCAT CACCATTAA  
        879

## 45                   (2) INFORMATION FOR SEQ ID NO:10:

- 50                   (i) SEQUENCE CHARACTERISTICS:  
                   (A) LENGTH: 293 amino acids  
                   (B) TYPE: amino acid  
                   (C) STRANDEDNESS: single  
                   (D) TOPOLOGY: linear  
                  CLYTA E6 His HPV 16

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

55                   Met Lys Gly Gly Ile Val His Ser Asp Gly Ser Tyr Pro Lys Asp Lys  
        1                   5                   10                   15  
                  Phe Glu Lys Ile Asn Gly Thr Trp Tyr Tyr Phe Asp Ser Ser Gly Tyr  
        20                   25                   30  
        60                   Met Leu Ala Asp Arg Trp Arg Lys His Thr Asp Gly Asn Trp Tyr Trp  
        35                   40                   45  
                  Phe Asp Asn Ser Gly Glu Met Ala Thr Gly Trp Lys Lys Ile Ala Asp  
        50                   55                   60  
        65                   Lys Trp Tyr Tyr Phe Asn Glu Glu Gly Ala Met Lys Thr Gly Trp Val  
                  70                   75                   80  
                  Lys Tyr Lys Asp Thr Trp Tyr Tyr Leu Asp Ala Lys Glu Gly Ala Met

	85	90	95
Val Ser Asn Ala Phe Ile Gln Ser Ala Asp Gly Thr Gly Trp Tyr Tyr			
100	105	110	
Leu Lys Pro Asp Gly Thr Leu Ala Asp Arg Pro Glu Leu Ala Ser Met			
115	120	125	
Leu Asp Met Ala Met Phe Gln Asp Pro Gln Glu Arg Pro Arg Lys Leu			
130	135	140	
Pro Gln Leu Cys Thr Glu Leu Gln Thr Thr Ile His Asp Ile Ile Leu			
145	150	155	160
Glu Cys Val Tyr Cys Lys Gln Gln Leu Leu Arg Arg Glu Val Tyr Asp			
165	170	175	
Phe Ala Phe Arg Asp Leu Cys Ile Val Tyr Arg Asp Gly Asn Pro Tyr			
180	185	190	
Ala Val Cys Asp Lys Cys Leu Lys Phe Tyr Ser Lys Ile Ser Glu Tyr			
195	200	205	
Arg His Tyr Cys Tyr Ser Leu Tyr Gly Thr Thr Leu Glu Gln Gln Tyr			
210	215	220	
Asn Lys Pro Leu Cys Asp Leu Leu Ile Arg Cys Ile Asn Cys Gln Lys			
225	230	235	240
Pro Leu Cys Pro Glu Glu Lys Gln Arg His Leu Asp Lys Lys Gln Arg			
245	250	255	
Phe His Asn Ile Arg Gly Arg Trp Thr Gly Arg Cys Met Ser Cys Cys			
260	265	270	
Arg Ser Ser Arg Thr Arg Arg Glu Thr Gln Leu Thr Ser Gly His His			
275	280	285	
His His His His			
290			

(2) INFORMATION FOR SEQ ID NO:11:

30

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 720 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(B) TOPOLOGY: LINE

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

40 ATGAAAGGGG GAATTGTACA TTCAGACGGC TCTTATCCAA AAGACAAGTT TGAGAAAATC  
60 AATGGCACTT GGTACTACTT TGACAGTTCA GGCTATATGC TTGCAGACCG CTGGAGGAAG  
120 CACACAGACG GCAACTGGTA CTGGTTGAC AACTCAGGGG AAATGGCTAC AGGCTGGAAG  
45 180 AAAATCGCTG ATAAGTGGTA CTATTCAAC GAAGAAGGTG CCATGAAGAC AGGCTGGTCA  
240 AAGTACAAGG ACACCTGGTA CTACTTAGAC GCTAAAGAAG GCGCCATGGT ATCAAATGCC  
300 360 TTTATCCAGT CAGCGGACGG AACAGGCTGG TACTACCTCA AACAGACGG AACACTGGCA  
GACAGGCCAG ATTGGCCAG CATGCTGGAC ATGGCCATGC ATGGAGATAAC CCTACATTG  
420 CATGAATATA TGTTAGATT GCAACCAGAG ACAACTGATC TCTACTGTTA TGAGCAATT  
55 480 AATGACAGCT CAGAGGAGGA GGATGAAATA GATGGTCCAG CTGGACAAGC AGAACCGGC  
540 AGAGCCCATT ACAATATTGT AACCTTTGT TGCAAGTGTG ACTCTACGCT TCGGTTGTGC  
60 600 GTACAAAGCA CACACGTAGA CATTGTA CTTGGAAGACC TGTTAATGGG CACACTAGGA  
660 ATTGTGTGCC CCATCTGTT TCAGAAACCA ACTAGTGGCC ACCATCACCA TCACCATTAA  
720

65 (2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 240 amino acids  
(B) TYPE: amino acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear  
CLYTA E7 HIS HPV 16

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

Met Lys Gly Gly Ile Val His Ser Asp Gly Ser Tyr Pro Lys Asp Lys  
1 5 10 15  
Phe Glu Lys Ile Asn Gly Thr Trp Tyr Tyr Phe Asp Ser Ser Gly Tyr  
20 25 30  
Met Leu Ala Asp Arg Trp Arg Lys His Thr Asp Gly Asn Trp Tyr Trp  
35 40 45  
Phe Asp Asn Ser Gly Glu Met Ala Thr Gly Trp Lys Lys Ile Ala Asp  
50 55 60  
Lys Trp Tyr Tyr Phe Asn Glu Glu Gly Ala Met Lys Thr Gly Trp Val  
65 70 75 80  
Lys Tyr Lys Asp Thr Trp Tyr Tyr Leu Asp Ala Lys Glu Gly Ala Met  
85 90 95  
Val Ser Asn Ala Phe Ile Gln Ser Ala Asp Gly Thr Gly Trp Tyr Tyr  
100 105 110  
Leu Lys Pro Asp Gly Thr Leu Ala Asp Arg Pro Glu Leu Ala Ser Met  
115 120 125  
Leu Asp Met Ala Met His Gly Asp Thr Pro Thr Leu His Glu Tyr Met  
130 135 140  
Leu Asp Leu Gln Pro Glu Thr Thr Asp Leu Tyr Cys Tyr Glu Gln Leu  
145 150 155 160  
Asn Asp Ser Ser Glu Glu Asp Glu Ile Asp Gly Pro Ala Gly Gln  
165 170 175  
Ala Glu Pro Asp Arg Ala His Tyr Asn Ile Val Thr Phe Cys Cys Lys  
180 185 190  
Cys Asp Ser Thr Leu Arg Leu Cys Val Gln Ser Thr His Val Asp Ile  
195 200 205  
Arg Thr Leu Glu Asp Leu Leu Met Gly Thr Leu Gly Ile Val Cys Pro  
210 215 220  
Ile Cys Ser Gln Lys Pro Thr Ser Gly His His His His His His  
225 230 235

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:  
(A) LENGTH: 1173 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: single  
(D) TOPOLOGY: linear  
CLYTA E6E7 His HPV16

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

ATGAAAGGGG GAATTGTACA TTCAAGACGGC TCTTATCCAA AAGACAAGTT TGAGAAAATC  
60  
AATGGCACTT GGTACTACTT TGACAGTTCA GGCTATATGC TTGCAGACCG CTGGAGGAAG  
45  
120  
CACACAGACG GCAACTGGTA CTGGTTCGAC AACTCAGGGCG AAATGGCTAC AGGCTGGAAAG  
180  
AAAATCGCTG ATAAGTGGTA CTATTTCAAC GAAGAAGGTG CCATGAAGAC AGGCTGGGT  
240  
AACTACAAGG ACACTTGGTA CTACTTAGAC GCTAAAGAAG GCGCCATGGT ATCAAATGCC  
300  
TTTATCCAGT CAGCGGACGG AACAGGCTGG TACTACCTCA AACCAAGACGG AACACTGGCA  
360  
GACAGGCCAG ATTGGCCAG CATGCTGGAC ATGGCCATGT TTCAGGGACCC ACAGGAGCGA  
420

CCCAGAAAGT TACCACAGTT ATGCACAGAG CTGCAAACAA CTATACATGA TATAATATTA  
 480  
 GAATGTGTGT ACTGCAAGCA ACAGTTACTG CGACGTGAGG TATATGACTT TGCTTTCGG  
 540  
 5 GATTATGCA TAGTATATAG AGATGGGAAT CCATATGCTG TATGTGATAA ATGTTAAAG  
 600  
 6 TTTTATTCTA AAATTAGTGA GTATAGACAT TATTGTTATA GTTGTATGG AACAAACATTA  
 660  
 7 GAACAGCAAT ACAACAAACC GTTGTGTGAT TTGTTAATTA GGTGTATTAA CTGTCAAAAG  
 10 720  
 8 CCACTGTGTC CTGAAGAAAA GCAAAGACAT CTGGACAAAA AGCAAAGATT CCATAATATA  
 780  
 9 AGGGTCGGT GGACCGGTG ATGTATGTCT TGTGAGAT CATCAAGAAC ACGTAGAGAA  
 840  
 15 ACCCAGCTGA TGCATGGAGA TACACCTACA TTGCATGAAT ATATGTTAGA TTTGCAACCA  
 900  
 10 GAGACAACGT ATCTCTACTG TTATGAGCAA TTAAATGACA GCTCAGAGGA GGAGGATGAA  
 960  
 15 ATAGATGGTC CAGCTGGACA AGCAGAACCG GACAGAGCCC ATTACAATAT TGTAACCTTT  
 20 1020  
 20 TGTTGCAAGT GTGACTCTAC GCTTCGGTTG TCGTACAAA GCACACACGT AGACATTGCT  
 1080  
 25 ACTTTGGAAG ACCTGTTAAT GGGCACACTA GGAATTGTGT GCCCCATCTG TTCTCAGAAA  
 1140  
 25 CCAACTAGTG GCCACCATCA CCATCACCAT TAA  
 1173

## (2) INFORMATION FOR SEQ ID NO:14:

30 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 391 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear  
 35 CLYTA E6E7 His HPV16

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

	Met	Lys	Gly	Gly	Ile	Val	His	Ser	Asp	Gly	Ser	Tyr	Pro	Lys	Asp	Lys
40	1				5					10				15		
	Phe	Glu	Lys	Ile	Asn	Gly	Thr	Trp	Tyr	Tyr	Phe	Asp	Ser	Ser	Gly	Tyr
							20		25					30		
	Met	Leu	Ala	Asp	Arg	Trp	Arg	Lys	His	Thr	Asp	Gly	Asn	Trp	Tyr	Trp
45							35		40				45			
	Phe	Asp	Asn	Ser	Gly	Glu	Met	Ala	Thr	Gly	Trp	Lys	Ile	Ala	Asp	
							50		55			60				
	Lys	Trp	Tyr	Tyr	Phe	Asn	Glu	Gly	Ala	Met	Lys	Thr	Gly	Trp	Val	
50							65		70		75		80			
	Lys	Tyr	Lys	Asp	Thr	Trp	Tyr	Tyr	Leu	Asp	Ala	Lys	Glu	Gly	Ala	Met
							85		90			95				
	Val	Ser	Asn	Ala	Phe	Ile	Gln	Ser	Ala	Asp	Gly	Thr	Gly	Trp	Tyr	Tyr
55							100		105			110				
	Leu	Lys	Pro	Asp	Gly	Thr	Leu	Ala	Asp	Arg	Pro	Glu	Leu	Ala	Ser	Met
							115		120			125				
	Leu	Asp	Met	Ala	Met	Phe	Gln	Asp	Pro	Gln	Glu	Arg	Pro	Arg	Lys	Leu
60							130		135			140				
	Pro	Gln	Leu	Cys	Thr	Glu	Leu	Gln	Thr	Thr	Ile	His	Asp	Ile	Ile	Leu
							145		150		155		160			
	Glu	Cys	Val	Tyr	Cys	Lys	Gln	Gln	Leu	Leu	Arg	Arg	Glu	Val	Tyr	Asp
65							165		170			175				
	Phe	Ala	Phe	Arg	Asp	Leu	Cys	Ile	Val	Tyr	Arg	Asp	Gly	Asn	Pro	Tyr
							180		185			190				
	Ala	Val	Cys	Asp	Lys	Cys	Leu	Lys	Phe	Tyr	Ser	Lys	Ile	Ser	Glu	Tyr
							195		200			205				
	Arg	His	Tyr	Cys	Tyr	Ser	Leu	Tyr	Gly	Thr	Thr	Leu	Glu	Gln	Gln	Tyr
65							210		215			220				

Asn Lys Pro Leu Cys Asp Leu Leu Ile Arg Cys Ile Asn Cys Gln Lys  
 225                   230                   235                   240  
 Pro Leu Cys Pro Glu Glu Lys Gln Arg His Leu Asp Lys Lys Gln Arg  
 245                   250                   255  
 5 Phe His Asn Ile Arg Gly Arg Trp Thr Gly Arg Cys Met Ser Cys Cys  
 260                   265                   270  
 Arg Ser Ser Arg Thr Arg Arg Glu Thr Gln Leu Met His Gly Asp Thr  
 275                   280                   285  
 Pro Thr Leu His Glu Tyr Met Leu Asp Leu Gln Pro Glu Thr Thr Asp  
 10 290                   295                   300  
 Leu Tyr Cys Tyr Glu Gln Leu Asn Asp Ser Ser Glu Glu Asp Glu  
 305                   310                   315                   320  
 Ile Asp Gly Pro Ala Gly Gln Ala Glu Pro Asp Arg Ala His Tyr Asn  
 325                   330                   335  
 15 Ile Val Thr Phe Cys Cys Lys Cys Asp Ser Thr Leu Arg Leu Cys Val  
 340                   345                   350  
 Gln Ser Thr His Val Asp Ile Arg Thr Leu Glu Asp Leu Leu Met Gly  
 355                   360                   365  
 Thr Leu Gly Ile Val Cys Pro Ile Cys Ser Gln Lys Pro Thr Ser Gly  
 20 370                   375                   380  
 His His His His His  
 385                   390

## (2) INFORMATION FOR SEQ ID NO:15:

- 25 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 684 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 30 (D) TOPOLOGY: linear  
 Protein D 1/3 E7 his HPV 18

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

35 ATGGATCCAA GCAGCCATTC ATCAAATATG GCGAATAACCC AAATGAAATC AGACAAAATC  
 60 ATTATTGCTC ACCGTGGTGC TAGCGTTAT TTACCAAGAGC ATACGTTAGA ATCTAAAGCA  
 120 CTTGCGTTG CACAACAGGC TGATTATTTA GAGCAAGATT TAGCAATGAC TAAGGATGGT  
 180 CGTTTAGTGG TTATTACAGA TCACTTTTA GATGGCTTGA CTGATGTTGC GAAAAAAATC  
 240 CCACATCGTC ATCGTAAAGA TGGCCGTTAC TATGTCATCG ACTTTACCTT AAAAGAAATT  
 45 300 CAAAGTTTAG AAATGACAGA AAACTTGAA ACCATGGCCA TGCATGGACC TAAGGCAACA  
 360 TTGCAAGACA TTGTATTGCA TTTAGAGCCC CAAAATGAAA TTCCGGTTGA CCTTCTATGT  
 420 300 CACGAGCAAT TAAGCGACTC AGAGGAAGAA AACGATGAAA TAGATGAAGT TAATCATCAA  
 480 CATTTACCAAG CCCGACGAGC CGAACACCAA CGTCACACAA TGTTGTGTAT GTGTTGTAAG  
 540 55 600 TGTGAAGCCA GAATTGAGCT AGTAGTAGAA AGCTCAGCAG ACGACCTTCG AGCATTCCAG  
 CAGCTGTTTC TGAACACCCCT GTCCTTGTG TGTCCGTGGT GTGCATCCCA GCAGACTAGT  
 660 684 GCCCACCATC ACCATCACCA TTAA

## (2) INFORMATION FOR SEQ ID NO:16:

- 60 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 228 amino acids  
 65 (B) TYPE: amino acid  
 (C) STRANDEDNESS: single

(D) TOPOLOGY: linear  
Protein D 1/3 E7 his HPV 18

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

5       Met Asp Pro Ser Ser His Ser Ser Asn Met Ala Asn Thr Gln Met Lys  
      1                                 5                                 10                             15  
      Ser Asp Lys Ile Ile Ile Ala His Arg Gly Ala Ser Gly Tyr Leu Pro  
      20                                     25                                 30  
10      Glu His Thr Leu Glu Ser Lys Ala Leu Ala Phe Ala Gln Gln Ala Asp  
      35                                     40                                 45  
      Tyr Leu Glu Gln Asp Leu Ala Met Thr Lys Asp Gly Arg Leu Val Val  
      50                                     55                                 60  
15      Ile His Asp His Phe Leu Asp Gly Leu Thr Asp Val Ala Lys Lys Phe  
      65                                     70                                 75                             80  
      Pro His Arg His Arg Lys Asp Gly Arg Tyr Tyr Val Ile Asp Phe Thr  
      85                                     90                                 95  
      Leu Lys Glu Ile Gln Ser Leu Glu Met Thr Glu Asn Phe Glu Thr Met  
      100                                    105                                 110  
20      Ala Met His Gly Pro Lys Ala Thr Leu Gln Asp Ile Val Leu His Leu  
      115                                    120                                 125  
      Glu Pro Gln Asn Glu Ile Pro Val Asp Leu Leu Cys His Glu Gln Leu  
      130                                    135                                 140  
25      Ser Asp Ser Glu Glu Glu Asn Asp Glu Ile Asp Glu Val Asn His Gln  
      145                                    150                                 155                             160  
      His Leu Pro Ala Arg Arg Ala Glu Pro Gln Arg His Thr Met Leu Cys  
      165                                    170                                 175  
      Met Cys Cys Lys Cys Glu Ala Arg Ile Glu Leu Val Val Glu Ser Ser  
      180                                    185                                 190  
30      Ala Asp Asp Leu Arg Ala Phe Gln Gln Leu Phe Leu Asn Thr Leu Ser  
      195                                    200                                 205  
      Phe Val Cys Pro Trp Cys Ala Ser Gln Gln Thr Ser Gly His His His  
      210                                    215                                 220  
      His His His  
35      225

## (2) INFORMATION FOR SEQ ID NO:17:

## (i) SEQUENCE CHARACTERISTICS:

- 40      (A) LENGTH: 110 amino acids  
       (B) TYPE: amino acid  
       (C) STRANDEDNESS: single  
       (D) TOPOLOGY: linear  
       Thioredoxin

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

Met Ser Asp Lys Ile Ile His Leu Thr Asp Asp Ser Phe Asp Thr Asp  
1   5                                     10                             15  
50      Val Leu Lys Ala Asp Gly Ala Ile Leu Val Asp Phe Trp Ala Glu Trp  
      20                                    25                                     30  
      Cys Gly Pro Cys Lys Met Ile Ala Pro Ile Leu Asp Glu Ile Ala Asp  
      35                                    40                                     45  
      Glu Tyr Gln Gly Lys Leu Thr Val Ala Lys Leu Asn Ile Asp Gln Asn  
      50                                    55                                     60  
55      Pro Gly Thr Ala Pro Lys Tyr Gly Ile Arg Gly Ile Pro Thr Leu Leu  
      65                                    70                                     75                             80  
      Leu Phe Lys Asn Gly Glu Val Ala Ala Thr Lys Val Gly Ala Leu Ser  
      85                                    90                                     95  
60      Lys Gly Gln Leu Lys Glu Phe Leu Asp Ala Asn Leu Ala  
      100                                    105

## (2) INFORMATION FOR SEQ ID NO:18:

- 65      (i) SEQUENCE CHARACTERISTICS:  
       (A) LENGTH: 684 base pairs

(B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear  
 Protein D 1/3 E7 mutated HPV 18

5 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

ATGGATCCAA GCAGCCATT C ATCAAATATG GCGAATACCC AAATGAAAATC AGACAAAATC  
 60 ATTATTGCTC ACCGTGGTGC TAGCGTTAT TTACCAAGAGC ATACGTTAGA ATCTAAAGCA  
 10 120 CTTGCGTTG CACAACAGGC TGATTATTTA GAGCAAGATT TAGCAATGAC TAAGGATGGT  
 180 CGTTTAGTGG TTATTACAGA TCACTTTTA GATGGCTTGA CTGATGTTGC GAAAAAAATTC  
 15 240 CCACATCGTC ATCGTAAAGA TGGCCGTTAC TATGTCATCG ACTTTACCTT AAAAGAAATT  
 300 CAAAGTTTAG AAATGACAGA AAACCTTGAA ACCATGGCCA TGCATGGACC TAAGGCAACA  
 360 20 420 TTGCAAGACA TTGTATTGCA TTTAGAGCCC CAAAATGAAA TTCCGGTTGA CCTTCTAGGT  
 480 CACCAAGCAAT TAAGCGACTC AGAGGAAGAA AACGATGAAA TAGATGGAGT TAATCATCAA  
 540 25 600 TGTGAAGCCA GAATTGAGCT AGTAGTAGAA AGCTCAGCAG ACGACCTTCG AGCATTCCAG  
 660 CAGCTGTTTC TGAACACCCCT GTCCTTGTTG TGTCCGTGGT GTGCATCCCA GCAGACTAGT  
 30 684 GGCCACCATC ACCATCACCA TTAA  
 684

(2) INFORMATION FOR SEQ ID NO:19:

35 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 228 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear  
 40 Protein D 1/3 E7 mutated HPV 18

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

45 Met Asp Pro Ser Ser His Ser Ser Asn Met Ala Asn Thr Gln Met Lys  
 1 5 10 15  
 Ser Asp Lys Ile Ile Ile Ala His Arg Gly Ala Ser Gly Tyr Leu Pro  
 20 25 30  
 Glu His Thr Leu Glu Ser Lys Ala Leu Ala Phe Ala Gln Gln Ala Asp  
 35 40 45  
 Tyr Leu Glu Gln Asp Leu Ala Met Thr Lys Asp Gly Arg Leu Val Val  
 50 55 60  
 Ile His Asp His Phe Leu Asp Gly Leu Thr Asp Val Ala Lys Lys Phe  
 65 70 75 80  
 Pro His Arg His Arg Lys Asp Gly Arg Tyr Tyr Val Ile Asp Phe Thr  
 85 90 95  
 Leu Lys Glu Ile Gln Ser Leu Glu Met Thr Glu Asn Phe Glu Thr Met  
 100 105 110  
 Ala Met His Gly Pro Lys Ala Thr Leu Gln Asp Ile Val Leu His Leu  
 115 120 125  
 Glu Pro Gln Asn Glu Ile Pro Val Asp Leu Leu Gly His Gln Gln Leu  
 130 135 140  
 Ser Asp Ser Glu Glu Glu Asn Asp Glu Ile Asp Gly Val Asn His Gln  
 145 150 155 160  
 His Leu Pro Ala Arg Arg Ala Glu Pro Gln Arg His Thr Met Leu Cys  
 165 170 175

Met Cys Cys Lys Cys Glu Ala Arg Ile Glu Leu Val Val Glu Ser Ser  
 180 185 190  
 Ala Asp Asp Leu Arg Ala Phe Gln Gln Leu Phe Leu Asn Thr Leu Ser  
 195 200 205  
 5 Phe Val Cys Pro Trp Cys Ala Ser Gln Gln Thr Ser Gly His His His  
 210 215 220  
 His His His  
 225

10 (2) INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 837 base pairs  
 (B) TYPE: nucleic acid  
 15 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear  
 Protein D 1/3 E6 - His HPV 18

20 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

ATGGATCCAA GCAGCCATTG ATCAAATATG GCGAATAACCC AAATGAAATC AGACAAAATC  
 60 ATTATTGCTC ACCGTGGTGC TAGCGTTAT TTACCAAGAGC ATACGTTAGA ATCTAAAGCA  
 120 CTTGCCTTG CACAACAGGC TGATTATTTA GAGCAAGATT TAGCAATGAC TAAGGATGGT  
 180 CGTTTAGTGG TTATTACACGA TCACTTTTA GATGGCTTGA CTGATGTTGC GAAAAAAATC  
 240 CCACATCGTC ATCGTAAAGA TGGCCGTTAC TATGTCATCG ACTTTACCTT AAAAGAAATT  
 300 CAAAGTTTAG AAATGACAGA AAACCTTGAA ACCATGGCGC GCTTTGAGGA TCCAACACGG  
 360 CGACCCCTACA AGCTACCTGA TCTGTGCACG GAACTGAACA CTTCACTGCA AGACATAGAA  
 420 480 ATAACCTGTG TATATTGCAA GACAGTATTG GAACTTACAG AGGTATTGAA ATTTGCATTT  
 540 AAAGATTAT TTGTGGTGTAA TAGAGACAGT ATACCGCATG CTGCATGCCA TAAATGTATA  
 540 GATTTTATT CTAGAATTAG AGAATTAAGA CATTATTCAG ACTCTGTGTAA TGGAGACACA  
 40 600 TTGGAAAAAC TAACTAACAC TGGGTTATAC AATTTATTAA TAAGGTGCCT GCGGTGCCAG  
 660 AAACCGTTGA ATCCAGCAGA AAAACTTAGA CACCTTAATG AAAAACGACG ATTCACAAAC  
 720 780 ATAGCTGGGC ACTATAGAGG CCAGTGCAT TCAGTGCATGCA ACCGAGCACG ACAGGAACGA  
 837 CTCCAACGAC GCAGAGAAC ACAAGTAAC AGTGGCCACC ATCACCATCA CCATTAA

50 (2) INFORMATION FOR SEQ ID NO:21:

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 279 amino acids  
 (B) TYPE: amino acid  
 55 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear  
 Protein D 1/3 E6 - His HPV 18

60 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

Met Asp Pro Ser Ser His Ser Ser Asn Met Ala Asn Thr Gln Met Lys  
 1 5 10 15  
 Ser Asp Lys Ile Ile Ile Ala His Arg Gly Ala Ser Gly Tyr Leu Pro  
 20 25 30  
 65 Glu His Thr Leu Glu Ser Lys Ala Leu Ala Phe Ala Gln Gln Ala Asp  
 35 40 45

Tyr Leu Glu Gln Asp Leu Ala Met Thr Lys Asp Gly Arg Leu Val Val  
 50 55 60  
 Ile His Asp His Phe Leu Asp Gly Leu Thr Asp Val Ala Lys Lys Phe  
 65 70 75 80  
 5 Pro His Arg His Arg Lys Asp Gly Arg Tyr Tyr Val Ile Asp Phe Thr  
 85 90 95  
 Leu Lys Glu Ile Gln Ser Leu Glu Met Thr Glu Asn Phe Glu Thr Met  
 100 105 110  
 Ala Arg Phe Glu Asp Pro Thr Arg Arg Pro Tyr Lys Leu Pro Asp Leu  
 115 120 125  
 10 Cys Thr Glu Leu Asn Thr Ser Leu Gln Asp Ile Glu Ile Thr Cys Val  
 130 135 140  
 Tyr Cys Lys Thr Val Leu Glu Leu Thr Glu Val Phe Glu Phe Ala Phe  
 145 150 155 160  
 15 Lys Asp Leu Phe Val Val Tyr Arg Asp Ser Ile Pro His Ala Ala Cys  
 165 170 175  
 His Lys Cys Ile Asp Phe Tyr Ser Arg Ile Arg Glu Leu Arg His Tyr  
 180 185 190  
 Ser Asp Ser Val Tyr Gly Asp Thr Leu Glu Lys Leu Thr Asn Thr Gly  
 20 195 200 205  
 Leu Tyr Asn Leu Leu Ile Arg Cys Leu Arg Cys Gln Lys Pro Leu Asn  
 210 215 220  
 Pro Ala Glu Lys Leu Arg His Leu Asn Glu Lys Arg Arg Phe His Asn  
 225 230 235 240  
 25 Ile Ala Gly His Tyr Arg Gly Gln Cys His Ser Cys Cys Asn Arg Ala  
 245 250 255  
 Arg Gln Glu Arg Leu Gln Arg Arg Glu Thr Gln Val Thr Ser Gly  
 260 265 270  
 His His His His His His  
 30 275

## (2) INFORMATION FOR SEQ ID NO:22:

- (i) SEQUENCE CHARACTERISTICS:  
 35 (A) LENGTH: 1152 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

40 Protein D1/3 E6 E7 His/ HPV 18

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

ATGGATCCAA GCAGCCATTC ATCAAATATG GCGAATACCC AAATGAAATC AGACAAAATC  
 60  
 45 ATTATTGCTC ACCGTGGTGC TAGCGTTAT TTACCAGAGC ATACGTTAGA ATCTAAAGCA  
 120 CTTGCGTTG CACAACAGGC TGATTATTTA GAGCAAGATT TAGCAATGAC TAAGGATGGT  
 180 CGTTTAGTGG TTATTCACGA TCACTTTTA GATGGCTTGA CTGATGTTGC GAAAAAAATC  
 50 240 CCACATCGTC ATCGTAAAGA TGGCCGTTAC TATGTCATCG ACTTTACCTT AAAAGAAATT  
 300 CAAAGTTTAG AAATGACAGA AAACTTGAA ACCATGGCCG GCTTGAGGA TCCAACACGG  
 360 55 CGACCCTACA AGCTACCTGA TCTGTGCACG GAACTAACCA CTTCACTGCA AGACATAGAA  
 420  
 480 ATAACCTGTG TATATTGCAA GACAGTATTG GAACTTACAG AGGTATTTGA ATTTGCATTT  
 540 AAAGATTTAT TTGTGGTGTAGAGACAGT ATACCGCATG CTGCATGCCA TAAATGTATA  
 60 GATTTTTATT CTAGAATTAG AGAATTAAGA CATTATTCAAG ACTCTGTGTA TGGAGACACA  
 660 TTGGAAAAAC TAACTAACAC TGGGTTATAC AATTTATTAA TAAGGTGCCT GCGGTGCCAG  
 660 AAACCGTTGA ATCCAGCAGA AAAACTTAGA CACCTTAATG AAAAACGACG ATTCACAAAC  
 720

ATAGCTGGGC ACTATAGAGG CCAGTGCCAT TCGTGCTGCA ACCGAGCACCG ACAGGAACGA  
 780  
 CTCCAACGAC GCAGAGAAC ACAAGTAATG CATGGACCTA AGGCAACATT GCAAGACATT  
 840  
 5 GTATTGCATT TAGAGCCCCA AAATGAAATT CCGGTTGACC TTCTATGTCA CGAGCAATTA  
 900  
 960 AGCGACTCAG AGGAAGAAAA CGATGAAATA GATGGAGTTA ATCATCAACA TTTACCAGCC  
 1020 CGACGAGCCG AACCACAACG TCACACAAATG TTGTGTATGT TTGTAAGTG TGAAGCCAGA  
 1080 ATTGAGCTAG TAGTAGAAAG CTCAGCAGAC GACCTTCGAG CATTCCAGCA GCTGTTCTG  
 1140 AACACCCCTGT CCTTTGTGTG TCCGTGGTGT GCATCCCAGC AGACTAGTGG CCACCATCAC  
 1152 15 CATCACCAATT AA  
 1152

## (2) INFORMATION FOR SEQ ID NO:23:

20 (i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 384 amino acids  
 (B) TYPE: amino acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear  
 25 Protein D1/3 E6 E7 His/ HPV 18

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

Met Asp Pro Ser Ser His Ser Ser Asn Met Ala Asn Thr Gln Met Lys  
 30 1 5 10 15  
 Ser Asp Lys Ile Ile Ile Ala His Arg Gly Ala Ser Gly Tyr Leu Pro  
 20 25 30  
 Glu His Thr Leu Glu Ser Lys Ala Leu Ala Phe Ala Gln Gln Ala Asp  
 35 40 45  
 35 Tyr Leu Glu Gln Asp Leu Ala Met Thr Lys Asp Gly Arg Leu Val Val  
 50 55 60  
 Ile His Asp His Phe Leu Asp Gly Leu Thr Asp Val Ala Lys Lys Phe  
 65 70 75 80  
 Pro His Arg His Arg Lys Asp Gly Arg Tyr Tyr Val Ile Asp Phe Thr  
 40 85 90 95  
 Leu Lys Glu Ile Gln Ser Leu Glu Met Thr Glu Asn Phe Glu Thr Met  
 100 105 110  
 Ala Arg Phe Glu Asp Pro Thr Arg Arg Pro Tyr Lys Leu Pro Asp Leu  
 115 120 125  
 45 Cys Thr Glu Leu Asn Thr Ser Leu Gln Asp Ile Glu Ile Thr Cys Val  
 130 135 140  
 Tyr Cys Lys Thr Val Leu Glu Leu Thr Glu Val Phe Glu Ala Phe  
 145 150 155 160  
 Lys Asp Leu Phe Val Val Tyr Arg Asp Ser Ile Pro His Ala Ala Cys  
 50 165 170 175  
 His Lys Cys Ile Asp Phe Tyr Ser Arg Ile Arg Glu Leu Arg His Tyr  
 180 185 190  
 Ser Asp Ser Val Tyr Gly Asp Thr Leu Glu Lys Leu Thr Asn Thr Gly  
 195 200 205  
 55 Leu Tyr Asn Leu Leu Ile Arg Cys Leu Arg Cys Gln Lys Pro Leu Asn  
 210 215 220  
 Pro Ala Glu Lys Leu Arg His Leu Asn Glu Lys Arg Arg Phe His Asn  
 225 230 235 240  
 Ile Ala Gly His Tyr Arg Gly Gln Cys His Ser Cys Cys Asn Arg Ala  
 60 245 250 255  
 Arg Gln Glu Arg Leu Gln Arg Arg Arg Glu Thr Gln Val Met His Gly  
 260 265 270  
 Pro Lys Ala Thr Leu Gln Asp Ile Val Leu His Leu Glu Pro Gln Asn  
 275 280 285  
 65 Glu Ile Pro Val Asp Leu Leu Cys His Glu Gln Leu Ser Asp Ser Glu  
 290 295 300

Glu Glu Asn Asp Glu Ile Asp Gly Val Asn His Gln His Leu Pro Ala  
305 310 315 320  
Arg Arg Ala Glu Pro Gln Arg His Thr Met Leu Cys Met Cys Cys Lys  
325 330 335  
5 Cys Glu Ala Arg Ile Glu Leu Val Val Glu Ser Ser Ala Asp Asp Leu  
340 345 350  
Arg Ala Phe Gln Gln Leu Phe Leu Asn Thr Leu Ser Phe Val Cys Pro  
355 360 365  
10 Trp Cys Ala Ser Gln Gln Thr Ser Gly His His His His His His  
370 375 380